Essays on International and Environmental Economics: A Game-theoretic Approach

Elboghdadly, Noha Nagi

Award date:
2020

Awarding institution:
University of Bath

Essays on International and Environmental Economics: A Game-theoretic Approach

submitted by

Noha Nagi Elboghdadly

for the degree of Doctor of Philosophy

of the

University of Bath

Department of Economics

December 2019

بسم الله الرحمن الرحيم

" رَّبُّ الْمَشْرِقِ وَالْمَغْرِبِ لَا إِلَٰهَ إِلَّا هُوَ فَاتَّخِذْهُ وَكِيلًا "

صدق الله العظيم

سورة المزمل الأيه 9

In the Name of Allah, the Most Beneficent, the Most Merciful

"He is Lord of the East and the West: there is no God but He, so take Him as Disposer of your affairs"

Quran 73: Verse 9

To my parents and sister

Acknowledgements

This work would not have been accomplished without the sustaining grace of Allah Almighty, as His blessings, inspiration, guidance and power have given me strength throughout my life.

Going through this PhD has been a rich and challenging experience. I would like to thank all the people who supported and guided me throughout this journey.

My special appreciation and gratitude go to my lead supervisor, Prof. Michael Finus. I appreciate his patience and continuous support from the very first day of when I started an entirely new field of research for me. He gave me invaluable comments and guidance that helped me to develop my ideas and research skills throughout my PhD studies. His stimulating discussions and the generous time he gave me even after he moved from Bath was more than I could hope for. He remained encouraging, kind, supportive and caring even during his holidays before my submission.

I am grateful to my second supervisor, Dr. Javier Rivas for his support and insightful comments that enlightened my ideas. I would also like to thank my third supervisor Dr. Paolo Zeppini for his helpful advice and support and for giving me the opportunity to gain valuable teaching experience with him. I appreciate all the helpful comments I received from the staff in the Department of Economics at the University of Bath during workshops and seminars. I would like to thank Dr. Maik Schneider and Dr. Ralph Winkler for the time they spent on reading my thesis, and for all their helpful and constructive comments.

I had the honour to meet beautiful souls and lovable friends in Bath. My special gratitude goes to my dear friend and office mate Maryam Al-Bulushi for being there for me with her deep understanding and lovable support that helped me a lot to maintain balance within my life. I would also like to thank Alaa, Bayan, Hanan, Mariam Essam, Marwa, Nahla, Rahmi, Sunday, Tayba, Van. We were as a family, advising, encouraging, helping and supporting each other.

I am grateful and indebted to the Egyptian Ministry of Higher Education which financially supported my PhD studies at the University of Bath.

Most importantly, none of this could have happened without my family that constitute the backbone of life: my father, my mother and my sister Nevine. No words can express my profound gratitude for their patience, support and faith in me throughout my whole life. From the very first day, they shared with me my dream and let me pursue my educational ambitions abroad despite their worries and the difficult feelings of my being away. They assisted and motivated me tirelessly throughout this journey. I am grateful for their efforts when they came occasionally to comfort and support me, and the accompany of my father at the beginning of my journey to Bath with his kindness and care. They are the pillar of my life, with their love and prayers have given me great strength during difficult times. I am praying to Allah to bless them and make them feel proud of me.

Noha Nagi Elboghdadly

December 2019

Contents

List of Figures

Thesis Abstract

Cooperative actions on climate change are difficult to achieve due to asymmetries among countries and free-rider incentives. In addition, the effectiveness of non-cooperative actions is undermined by carbon leakage. Using a game-theoretic analysis, the aim of this thesis is to study and evaluate the impact of different measures to mitigate climate change. The first essay in the thesis discusses the main features of the interaction among countries in mitigating climate change. We explain the trade-off between individual rationality and efficiency and demonstrate ways in which asymmetries among countries could affect the outcomes of climate negotiations.

The second essay investigates the incentives of governments when designing their non-cooperative climate policies under different policy regimes. We study the effect of a gradual shift from bilateral production-based carbon taxes to unilateral or bilateral consumption-based ones, considering various forms of trade measures called border carbon adjustments (BCAs). We find that although profit-shifting and carbon leakage distortions are only eliminated by combining carbon tariffs with a full export rebate, the optimal tax may still be below individual marginal damages. In contrast, a bilateral consumption-based tax could be set equal to or even above individual marginal damages.

The third essay investigates the conditions under which a sequence of escalating penalties of BCA-measures could be successful in enforcing a fully cooperative agreement. We show that import tariffs are the least distortionary policy instrument but the weakest punishment, and import tariffs with a full export rebate is the most distortionary instrument if implemented but the harshest punishment to enforce cooperation. However, whenever full cooperation is be expected to generate the highest global welfare gains, the harshest punishment fails to establish cooperation.

The fourth essay analyses the role of BCAs in a setting where the location of firms is chosen endogenously and countries choose their carbon taxes simultaneously or sequentially. We find that without BCAs, a 'race to the bottom' is the Nash equilibrium. In a Stackelberg equilibrium, a second less negative 'chicken equilibrium' may emerge. With BCAs, the race-to-the-bottom in carbon taxes can be avoided in the Nash equilibrium. However, a Nash equilibrium may not exist due to the discontinuity of best response functions. BCAs always reduce global emissions and in most cases increase global welfare under sequential choices of taxes.

Part I

Thesis Introduction

1 Background

International environmental problems, such as the climate change, have become important issues, both at the economic and political levels. The international aspect of these problems leads to considerable interdependence among countries. For instance, all countries contribute to and are affected by greenhouse gas (GHG) emissions, the major driver of global warming. Therefore, an effective solution to climate change requires cooperation among all countries in designing international instead of national climate policies to internalise global damages from emissions.

Negotiations to coordinate actions of nations on climate change started when the United Nations Framework Convention on Climate Change (UNFCCC) was adopted in 1992. The first historic binding agreement on climate change is the Kyoto Protocol, which came into force in 2005 with only 37 of the developed and industrialised countries agreeing to undertake binding targets to reduce their emissions. The effectiveness of the Kyoto Agreement has been questioned both due to the exemption of all developing countries including China and India from binding targets and the withdrawal of the United States (Böhringer and Finus, 2005; Nordhaus and Boyer, 1999). The Paris Agreement, which was adopted in 2015, is considered the second historic agreement to mitigate climate change, where both developed and developing countries agreed to reduce their GHG emissions (UNFCCC, 2015). However, the Paris Agreement is not binding. That is, countries have put forward their voluntary pledges in the form of nationally determined contributions. These pledges collectively will fall short of meeting the intended target of restraining global warming to below 2°C (Höhne et al., 2017). Moreover, the United States announced its intention to withdraw from the agreement, suggesting that a world of sub-global climate actions will still prevail, at least in the short run.

As reviewed above, the long history of climate negotiations reflects the difficulty of reaching a climate agreement. A wide strand of the game-theoretic literature analyses the challenges of cooperation and the formation of international environmental agreements (IEAs)(e.g., Barrett, 1994; Carraro and Siniscalco, 1993; Finus, 2008). This literature explains that in the absence of an enforcement power to make climate agreements legally binding, the cooperation of countries must be voluntary and must satisfy each country's self-interest. However, because it is a global public good, emission abatement is non-excludable, and thus countries have incentives to free-ride on the abatement efforts undertaken by other countries. In addition, countries are asymmetric regarding different political priorities, income levels or technologies. Those asymmetries among nations are also reflected in their valuation of environmental damages and willingness to abate emissions rendering cooperation not profitable for environmentally less concerned countries.

Early IEA literature shows that due to the above mentioned challenges, in particular free-rider incentives, the gains achieved from cooperation are not meaningful (Barrett, 1994). That is, either small stable agreements emerge or if there are large agreements, then the need for cooperation in terms of global welfare is not large. In an attempt to achieve more optimistic outcomes, the literature has been extended in different directions.[1] For instance, instead of assuming a cost effective allocation of abatement burdens among cooperating countries, some papers consider different bargaining rules such as emission reduction quotas (e.g., Altamirano-Cabrera et al., 2008; Endres and Finus, 2002; Finus and Rundshagen, 1998; Hoel, 1992). Some other papers introduce side payments, taking the form of different transfer schemes (e.g., Barrett, 2001; Botteon and Carraro, 1997; Carraro et al., 2006). Transfers, either monetary or in-kind, is an important 'carrot' instrument to enhance cooperation among asymmetric countries. However, monetary transfers are rarely observed in reality due to challenges of commitment among donors and receivers and between donors themselves (Barrett, 1994; Finus, 2002). In contrast, trade sanctions are considered as a 'stick' instrument to enforce cooperation (Barrett, 1997). A well known example of trade sanctions is prohibiting trade in chlorofluorocarbons products between signatories and non-signatories of the Montreal Protocol which was signed in 1987. However, trade sanctions face certain constraints as they should be credible but also compatible with international trade agreements.

Beside the interaction among countries in mitigating climate change, introducing other sources of interdependence among nations may add to the challenges of reducing global emissions. Two important sources are international trade and factor mobility, both of which are closely linked to climate change. Emission reductions by some environmentally friendly countries can be partly or completely offset by higher emission levels in other countries, a phenomenon known as 'carbon leakage'. This could occur through the relocation of the production of firms either in terms of market shares or the entire production facilities, to countries with less strict (without) climate policies. Carbon leakage raises concerns of countries about the effectiveness of their unilateral or sub-global actions, but also about the competitiveness of their firms particularly in emission-intensive trade-exposed (EITE) industries.

In order to address carbon leakage and protect domestic industries, border carbon adjustments (BCAs) have been recently proposed both by policy makers

[1] See Finus and Caparros (2015) for a recent survey of the literature.

(Baker et al., 2017; Davenport, 2016), and economists (e.g., Böhringer et al., 2012; Branger and Quirion, 2014; Fischer and Fox, 2012). BCAs are trade measures, adjusting the difference between the carbon prices of countries, and thus would be imposed by relatively more regulated countries. These border adjustment measures might be imposed on imports, taking the form of carbon tariffs, or on exports, taking the form of export rebates, or on imports and exports by combining both measures.

Additional arguments advocating BCAs suggest that they could incentivise countries on which these measures are imposed to implement comparable carbon policies and hence might assist in fostering climate agreements (Helm et al., 2012; Stiglitz, 2006). Furthermore, BCAs are also relevant in the debate concerning the allocation of emissions mitigation costs between consumers and producers. That is, complementing national climate policies with BCAs implies de facto regulating emissions based on consumption rather than production. Hence, some economists view BCAs as a way of switching to a partial or full consumption-based carbon price which might be more favourable than the current production-based approach (Peters and Hertwich, 2008; Steininger et al., 2014). Nevertheless, being unilateral trade measures, BCAs should be designed to meet the rules of the World Trade Organisation (WTO)(Horn and Mavroidis, 2011; Ismer and Neuhoff, 2007).

2 Thesis Context and Objectives

Against the above background, this thesis studies the strategic interaction among asymmetric countries facing the problem of climate change. The thesis consists of four self-contained essays which together contribute to the literature by analysing different measures to stimulate the efforts of nations to mitigate global emissions in cooperative or non-cooperative settings. Throughout the thesis, we consider two-country models under different levels of interdependence among countries.

Essay 1 explains the main features of the interaction among countries in mitigating climate change. In order to reach a fully cooperative outcome, cooperation must be individually rational (profitable). In the absence of transfers, a socially optimal solution may not be achieved if countries are sufficiently asymmetric in terms of their benefits and costs of emission abatement. We show how the type and degree of asymmetry among countries would affect the outcomes of climate negotiations. In addition, we briefly discuss the impacts of some domestic policies on the negotiations like investment and adaptation.

Trade measures, like BCAs, could also affect climate negotiations. In order to analyse the effects of these measures in detail, the remaining three essays are devoted to studying the impacts of BCAs on the non-cooperative outcomes (Essay 2 and 4) and on enforcing a fully cooperative outcome (Essay 3). For this, we employ a strategic trade model, which is an extended version of Brander and Spencer (1985). Essays 2, 3 and 4 consider two countries which evaluate the damages from global emissions differently. Both countries impose a carbon tax that affects not only their individual damages but also their governmental revenues, consumers, and the profits of firms which compete in a Cournot-fashion. As mentioned above, introducing international trade and factor mobility implies that countries do not confront only challenges of profitability or free-riding, but also competitiveness and carbon leakage issues, which we cover in these essays.

Essays 2 and 3 study the implications of a gradual shift from a production-based to consumption-based carbon tax regimes using BCA-measures, including carbon tariffs and export rebates. In Essay 2, we solve a two-stage game in which countries simultaneously choose their carbon taxes in the first stage, and then firms choose their output levels in the second stage. The focus of Essay 2 is to study the impacts of BCAs on the incentives of governments when designing their non-cooperative optimal climate policies. Whereas the focus of Essay 3 is to analyse the strategic role of different BCA-measures in enforcing a fully cooperative agreement and to evaluate their impacts on global emissions and global welfare if implemented. Thus, we extend the game in Essay 2 to three stages, such that countries decide whether to fully cooperate and implement a uniform socially optimal tax level in the first stage.

We assume that the location of firms is fixed in Essays 2 and 3. That is, these two essays consider carbon leakage due to the relocation of market shares. In Essay 4, we relax this assumption and allow for firm mobility, according to which, firms could relocate their entire production facilities to countries with lax climate policies. Here, we also extend the game of Essay 2 to three stages in which governments first choose their carbon tax, and then firms choose their location in the second stage and their output levels in the third stage. In this way, the focus of Essay 4 is to study the impacts of BCAs (carbon tariffs) on climate policy levels, global emissions and global welfare when firm locations are endogenous.

3 Summary of Essays

In this section, I will provide a brief summary of the contents and findings of each essay.

Essay 1: Negotiations on Climate Change Mitigation among Asymmetric Countries

Essay 1 presents the basic features of the climate change negotiations among asymmetric countries. A Pareto-efficient solution can be reached through maximising the weighted sum of the welfare functions of all countries. Depending on the weights attached to countries in the global welfare function, many cooperative outcomes could emerge from the negotiations. We compare the first-best (FB) solution, which follows from unconstrained joint welfare maximisation with the constrained joint welfare maximisation and the Nash bargaining solution (NBS). The type of asymmetry between benefits and costs associated with abatement greatly affects the trade-offs between efficiency and individual rationality. We provide some insights about the role of asymmetries in allocating abatement efforts and on global abatement under different Pareto-optimal outcomes. Finally, we show how transfers and domestic policies could change the outcomes of climate negotiations.

Essay 2: Non-Cooperative Climate Policies among Asymmetric Countries: Production- versus Consumption-based Carbon Taxes

Essay 2 studies the effect of a gradual shift from bilateral production- to unilateral or bilateral consumption-based carbon tax. We consider five non-cooperative regimes: a bilateral production-based tax (PB-regime) and a bilateral consumption-based tax (CB-regime) regime as well as three border carbon adjustments (BCAs) regimes, under which the environmentally more concerned country complements its carbon tax with carbon tariffs and export rebates. We assume that carbon tariffs fully adjust the difference between national carbon taxes, while two forms of export rebates are considered: optimal and full rebates. This essay contributes to the literature by analysing whether a switch to either a unilateral or bilateral consumption-based tax regime could restore the effectiveness of non-cooperative climate policies in a strategic context.

We find that carbon tariffs create a new incentive for both governments to tax emissions. In addition, combining carbon tariffs with full export rebate would eliminate the profit-shifting and carbon leakage effects. However, the effectiveness of carbon taxes of both countries could be restored to fully internalise individual damages only under the CB-regime. Our results show that adding export rebates reinforces (weakens) the effect of carbon tariffs on the climate policy level of the environmentally more (less) concerned country.

Essay 3: Enforcing Climate Agreements: The Role of Escalating Border Carbon Adjustments

Essays 3 considers the same BCA-regimes as Essay 2. We start our analysis from the assumption that moving from non-cooperative and non-uniform production-based carbon taxes to a uniform socially optimal tax is not attractive to the environmentally less concerned country. The decision of countries on cooperation is then based on a multi-stage escalating penalty game. In each stage, the environmentally more concerned country moves first and decides whether to use a BCA-threat to propose 'cooperation'. After this, the other country chooses whether to accept the proposal. We contribute to the literature by analysing whether and under which conditions a sequence of BCA-threats, which gradually lead to a unilateral consumption-based carbon tax, would enforce full cooperation. In addition, we evaluate the outcome in terms of global emissions and global welfare.

We find that our escalating sequence of BCA penalties is effective, but also credible to enforce full cooperation if environmental damages are not too large compared to the net benefits from production and consumption. However, this implies in our model that the potential gains from full cooperation would also not be too large. We show that if BCAs fail to establish cooperation, they lead to higher global welfare and lower global emission levels if they need to be implemented. However, the harsher a BCA-threat, the more distortionary it would be if implemented compared to the social optimum.

Essay 4: Strategic Climate Policies with Endogenous Plant Location: The Role of Border Carbon Adjustments

Essay 4 studies a carbon tax competition game between two governments that strive to attract the plants of two firms. We consider two regimes. Under the No-BCA regime, each government imposes a carbon tax on the production of plants located within its national boundaries. Under the BCA regime, the country that sets a higher carbon tax can additionally impose a carbon tariff on imports from plants located abroad. For each policy regime, we determine climate policy equilibria in a simultaneous and a sequential game. This essay contributes to the literature by investigating the impacts of BCAs on imports if governments consider the endogenous location choices of firms.

On the one hand, all plants will locate in the country which sets a lower carbon tax under a No-BCA regime. This leads to a fierce tax competition under the simultaneous choice of taxes. As a result, the Nash equilibrium is 'race to the bottom'. We show that moving sequentially could be Pareto-improving for both countries and lead to lower global emissions if the marginal damage of the Stackelberg leader is sufficiently high. On the other hand, the country which imposes BCAs is able to partially protect its home firm which supplies the domestic market. However, we find that a Nash equilibrium may not exist under the BCA regime. Nevertheless, if a Nash exists, BCAs lead to a more ambitious climate policy level in both countries and higher global welfare than under the No-BCA regime. Furthermore, the race-to-the-bottom equilibrium under the sequential choice of taxes is less likely to emerge with BCAs implying lower global emissions and in most cases higher global welfare levels.

References

Altamirano-Cabrera, J.-C., Finus, M., and Dellink, R. (2008). Do abatement quotas lead to a more successful climate coalition? *The Manchester School*, 51(1):93–109.

Baker, J. A. I., Feldstein, M., Halstead, T., Mankiw, N. G., Paulson Jr, H. M., Shultz, G. P., Stephenson, T., and Walton, R. (2017). The conservative case for carbon dividends. *Climate Leadership Council*.

Barrett, S. (1994). Self-enforcing international environmental agreements. *Oxford Economic Papers*, 46:878–894.

Barrett, S. (1997). The strategy of trade sanctions in international environmental agreements. *Resource and Energy Economics*, 19(4):345–361.

Barrett, S. (2001). International cooperation for sale. *European Economic Review*, 45(10):1835–1850.

Böhringer, C., Balistreri, E. J., and Rutherford, T. F. (2012). The role of border carbon adjustment in unilateral climate policy: overview of an energy modeling forum study EMF 29. *Energy Economics*, 34:S97–S110.

Böhringer, C. and Finus, M. (2005). The Kyoto Protocol: success or failure? In Helm, D., ed., *Climate Change Policy*, pp. 253–281, Oxford, UK. Oxford University Press.

Botteon, M. and Carraro, C. (1997). Burden-sharing and coalition stability in environmental negotiations with asymmetric countries. In Carraro, C., ed., *International Environmental Negotiations: Strategic Policy Issues*, pp. 26–55. Cheltenham, UK: Elgar.

Brander, J. A. and Spencer, B. J. (1985). Export subsidies and international market share rivalry. *Journal of International Economics*, 18(1-2):83–100.

Branger, F. and Quirion, P. (2014). Would border carbon adjustments prevent carbon leakage and heavy industry competitiveness losses? Insights from a meta-analysis of recent economic studies. *Ecological Economics*, 99:29–39.

Carraro, C., Eyckmans, J., and Finus, M. (2006). Optimal transfers and participation decisions in international environmental agreements. *The Review of International Organizations*, 1(4):379–396.

Carraro, C. and Siniscalco, D. (1993). Strategies for the international protection of the environment. *Journal of Public Economics*, 52(3):309–328.

Davenport, C. (2016). Diplomats confront new threat to Paris climate pact: Donald Trump. *New York Times*, 18.

Endres, A. and Finus, M. (2002). Quotas may beat taxes in a global emission game. *International Tax and Public Finance*, 9(6):687–707.

Finus, M. (2002). Game theory and international environmental cooperation: any practical application? In *Controlling Global Warming: Perspectives from Economics, Game Theory and Public Choice*, pp. 9–104. Edward Elgar Publishing Ltd.

Finus, M. (2008). Game theoretic research on the design of international environmental agreements: insights, critical remarks, and future challenges. *International Review of Environmental and Resource Economics*, 2(1):29–67.

Finus, M. and Caparros, A. (2015). *Game Theory and International Environmental Cooperation: Essential Readings*. The International Library of Critical Writings in Economics Series. Edward Elgar Publishing Ltd, UK United Kingdom.

Finus, M. and Rundshagen, B. (1998). Toward a positive theory of coalition formation and endogenous instrumental choice in global pollution control. *Public Choice*, 96(1-2):145–186.

Fischer, C. and Fox, A. K. (2012). Comparing policies to combat emissions leakage: border carbon adjustments versus rebates. *Journal of Environmental Economics and Management*, 64(2):199–216.

Helm, D., Hepburn, C., and Ruta, G. (2012). Trade, climate change, and the political game theory of border carbon adjustments. *Oxford Review of Economic Policy*, 28(2):368–394.

Hoel, M. (1992). International environment conventions: the case of uniform reductions of emissions. *Environmental and Resource Economics*, 2(2):141–159.

Höhne, N., Kuramochi, T., Warnecke, C., Röser, F., Fekete, H., Hagemann, M., Day, T., Tewari, R., Kurdziel, M., Sterl, S., et al. (2017). The Paris agreement: resolving the inconsistency between global goals and national contributions. *Climate Policy*, 17(1):16–32.

Horn, H. and Mavroidis, P. C. (2011). To b (ta) or not to b (ta)? On the legality and desirability of border tax adjustments from a trade perspective. *The World Economy*, 34(11):1911–1937.

Ismer, R. and Neuhoff, K. (2007). Border tax adjustment: a feasible way to support stringent emission trading. *European Journal of Law and Economics*, 24(2):137–164.

Nordhaus, W. D. and Boyer, J. G. (1999). Requiem for Kyoto: an economic analysis of the Kyoto Protocol. *The Energy Journal*, 20:93–130.

Peters, G. P. and Hertwich, E. G. (2008). Post-Kyoto greenhouse gas inventories: production versus consumption. *Climatic Change*, 86(1-2):51–66.

Steininger, K., Lininger, C., Droege, S., Roser, D., Tomlinson, L., and Meyer, L. (2014). Justice and cost effectiveness of consumption-based versus production-based approaches in the case of unilateral climate policies. *Global Environmental Change*, 24:75–87.

Stiglitz, J. (2006). A new agenda for global warming. *The Economists' Voice*, 3(7).

UNFCCC (2015). Report of the conference of the parties on its twenty-first session, held in Paris from 30 november to 13 december 2015. In *United Nations Framework Convention on Climate Change*.

Part II

Essay 1: Negotiations on Climate Change Mitigation among Asymmetric Countries

Appendix 6B: Statement of Authorship

This declaration concerns the article entitled:	
Negotiations on Climate Change Mitigation among Asymmetric Countries	
Publication status (tick one)	
Draft manuscript ☒ **Submitted** ☐ **In review** ☐ **Accepted** ☐ **Published** ☐	
Publication details (reference)	
Copyright status (tick the appropriate statement)	
I hold the copyright for this material ☒	Copyright is retained by the publisher, but I have been given permission to replicate the material here ☐
Candidate's contribution to the paper (provide details, and also indicate as a percentage)	The candidate contributed to / considerably contributed to / predominantly executed the… Formulation of ideas: - Considerably contributed to the formulation of ideas. (60 %) Design of methodology: - Predominantly contributed to the design of methodology. (70 %) ~~Experimental work:~~ Presentation of data in journal format: - Predominantly contributed to the presentation of data in journal format. (80%)
Statement from Candidate	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature.
Signed	Noha Nagi Elboghdadly **Date** 4/10/2019

Last update: Feb 2019

Negotiations on Climate Change Mitigation among Asymmetric Countries

Noha Elboghdadly* and Michael Finus†

Abstract

Solving climate change, the most notable example of international environmental problems, requires cooperation among countries to internalise global benefits from greenhouse gas abatement. The extensive history of climate change negotiations reflects the difficulty of achieving cooperation. In the absence of an enforcement power to make climate agreements legally binding, cooperation of countries must be voluntary. For this reason, individual rationality is a necessary condition to reach a cooperative outcome. This paper discusses the main features of the interaction among countries in mitigating climate change using game-theoretic analysis. The main focus of the paper is to analyse the effects of the asymmetry among countries in terms of their benefits and costs associated with climate change mitigation. Due to asymmetries among countries, there might be a trade-off between individual rationality and efficiency. We show how the type and degree of asymmetries could affect the negotiation outcomes in terms of the allocation of abatement burdens, global abatement and global welfare. We also discuss the effects of compensation measures such as transfers, and domestic policies such as investment and trade, on the outcomes of climate negotiations.

Keywords: Climate Change, Asymmetry, Pareto-optimal Outcomes, Nash Bargaining Solution.

JEL-Classification: C72, C78, H4, H87, Q54.

*Department of Economics, University of Bath, 3 East, Bath, BA2 7AY, UK. Email: n.m.w.elboghdadly@bath.ac.uk

†Department of Economics, University of Graz, Universitätsstraße 15, 8010 Graz, Austria and University of Bath, 3 East, Bath, BA2 7AY, UK. Email: michael.finus@uni-graz.at

1 Introduction

Climate change has become one of the most pressing environmental issues at the economic and political levels. The global nature of this problem implies that it affects all countries, irrespective of the location of emissions. This implies that climate change mitigation is a global public good that benefits all countries. Non-cooperative climate policies aiming to reduce greenhouse gas (GHG) emissions only consider a country's own benefits while the benefits to other countries are not internalised. This leads to the underprovision of the total abatement level required from the global point of view. Therefore, cooperation among countries is needed to reach a globally efficient outcome.

Despite the long history of climate change negotiations, which started in 1992 with the adoption of the United Nations Framework Convention on Climate Change (UNFCCC), only two climate agreements came into force, which reflects the difficulty of achieving cooperation. The first and only legally binding agreement is the Kyoto Protocol, which commits 37 industrialised countries to reduce their emissions. However, major global emitters such as China and the United States (USA) together with developing countries are not part of this agreement. The second one is the Paris Agreement, under which 190 countries have determined their voluntary pledges (nationally determined contributions) of emission reductions. However, those contributions will fall short of the 2-degree global warming target (Höhne et al., 2017).

The strategic interaction among countries in the process of climate change negotiations has attracted the attention of many economists and game-theorists. This paper provides an overview of the main features of the climate change negotiations using game theoretical models and focus on the role of asymmetry among countries with respect to their benefits and costs of GHG abatement. We are mainly interested to show how the type of asymmetry would affect the allocation of abatement burden between countries and the global abatement level under different Pareto-optimal solutions.

From a theoretical perspective, a non-cooperative solution to a global pollution problem is not Pareto-efficient. Hence, a Pareto-efficient outcome can be achieved through maximising a global welfare function, which is the weighted sum of the welfare functions of all countries (e.g., Chichilnisky and Heal, 1994; Escapa and Gutiérrez, 1997; Eyckmans et al., 1993; Hoel, 1991). Depending on the weights attached to countries in the global welfare function, many cooperative outcomes could emerge from negotiations.

First, we illustrate different Pareto-optimal solutions and the corresponding relative weights attached to countries in a global welfare function. The first-best (FB) solution, which follows from unconstrained joint welfare maximisation, implies equal weights attached to countries' welfare function. This, in turn, implies a cost-effective provision of abatement, where the marginal abatement costs of countries are equalised. Choosing abatement levels cost-effectively requires countries with relatively flatter marginal abatement cost curves to contribute more. However, if they perceive the benefits from joint abatement to be not very high, these countries might be made worse off under cooperation than under a non-cooperative outcome. That is, for those countries, cooperation may not be individually rational. Hence, the FB solution could not be achieved without transfers. For this, we consider two second-best (SB) solutions which are individually rational: the constrained joint welfare maximisation and the Nash bargaining solution (NBS).

Second, we shed more light on the role of asymmetry among countries in the climate negotiations. We start from the assumption that the FB solution may not be achieved due to asymmetries, and illustrate the trade-off between individual rationality and efficiency. The constrained joint welfare maximisation outcome is associated with the minimum global welfare loss compared to the FB solution, however, all gains from cooperation are captured by the country with a non-binding constraint. In contrast, the NBS leads to a more symmetric distribution of the gains from cooperation along with a higher global welfare loss than the constrained outcome. We then provide insights regarding ways in which the type of asymmetry affects the allocation of abatement burdens and global abatement. While most of the literature, e.g. Hoel (1992) and Boom (2006), that analyses the allocation of abatement between countries under the SB solutions focuses on one-sided asymmetry, i.e. either on the benefit or the cost side, we extend the analysis to different types of two-sided asymmetry. In addition, we show the impacts of these types of asymmetry on global abatement and on the performance of the SB solutions in closing the gap between no cooperation and the social optimum. We show that only if the country for which the individual rationality constraint is not binding has a higher marginal benefit and a flatter marginal cost curve, global abatement could be higher under the SB solutions compared to the FB solution. Though higher global abatement is also associated with a global welfare loss.

Finally, we briefly discuss certain domestic policies such as investment, adaptation and trade, which have been shown in the literature to affect the climate negotiation outcomes.

The early literature on environmental cooperation, in particular the formation of international environmental agreements (IEAs) assumes countries are symmetric (e.g., Barrett, 1994; Carraro and Siniscalco, 1993), and hence attaching equal weights to countries is innocuous. These papers conclude that only a small number of countries choose to cooperate due to strong free-rider incentives associated with the positive externality property of the IEAs. Although assuming symmetric countries is commonly used in the literature for the sake of simplicity, symmetry is far from reality.

Countries differ according to their stage of economic development, endowment of resources, population and level of technology. These disparities among nations are also reflected in their valuation of emission reduction. Therefore, many authors have relaxed this assumption, and alternatively have started to reflect the asymmetry among countries in their models (e.g., Barrett, 1997; Finus and McGinty, 2019; Hoel, 1992; Pavlova and De Zeeuw, 2013). However, if countries are heterogeneous, unconstrained joint welfare maximisation (i.e. attaching equal weights to countries) may result in an asymmetric distribution of the gains achieved from cooperation. To overcome this problem, many papers, for instance Barrett (2001); Botteon and Carraro (1997) and Carraro et al. (2006), consider different transfer schemes to compensate losers from cooperation.[1] Although transfers were found to be a vehicle for enhancing cooperation, they are rarely observed in reality. However, a few recent studies show that the type and degree of asymmetry among countries play a prominent role in the formation of IEAs, where asymmetries could complement or even substitute transfers (Finus and McGinty, 2019; McGinty, 2007; Pavlova and De Zeeuw, 2013).

In the absence of transfers, some literature departs from the FB solution and assumes bargaining rules to allocate the abatement duties among countries, which are referred to as second-best (SB) solutions. Under the SB solutions, countries are not treated equally, and their relative weights in a global welfare function are different. For instance, Escapa and Gutiérrez (1997), in a dynamic model, compare three cooperative outcomes: the FB solution, the Nash bargaining Solution (NBS) and the Kalai-Smorodinsky solution. Their main focus is to calculate the endogenous welfare weights corresponding to the three solutions for six countries.[2] They find that the weights countries receive in the global welfare function depend negatively on their gains from cooperation under the FB outcome. As a result,

[1] There are four transfer schemes which have been used in the IEAs literature. Three of them are related to concepts from cooperative game theory: the Nash bargaining solution, the Shapely value and the Chander-Tulkens transfer scheme, while the optimal transfer scheme is related to the non-cooperative game theory.

[2] See also Eyckmans et al. (1993) who compute the welfare weights of 12 regions for different designs of IEAs.

the highest weights in their simulations correspond to China and the USA which have flatter marginal cost functions relative to other regions.

While Pareto-optimal solutions are viewed as a normative approach, some other studies consider a positive analysis and assume that countries reduce their emissions by equal percentages from a certain base year. From the early papers, Hoel (1992) assumes that countries differ in their environmental damage but have the same abatement costs. He compares the FB solution with two cooperative outcomes: the constrained joint welfare maximisation, implying differentiated abatement levels, and a uniform emission reduction quota. He shows that the quota system is cost-inefficient and results in a higher global emission level and a lower global welfare than the FB outcome, while the global welfare loss under the constrained outcome is very small. Finus and Rundshagen (1998) compare two allocation rules; the uniform emission reduction quota and uniform emission tax. Although the latter is considered to be a cost-effective instrument, countries prefer the uniform quota over a uniform tax. A similar result has been reached by Endres and Finus (2002), namely that in equilibrium, an emission quota may achieve lower global emissions and a higher global welfare than an emission tax. Using the generalized Nash bargaining solution, Bayramoglu and Jacques (2015) compare the relative efficiency of two SB agreements: a uniform agreement with transfers and a differentiated agreement without transfers. Their analysis focuses on the effect of transfer payment costs on comparing the two agreements. They show that the first agreement is superior to the second one in terms of global and individual welfare levels if the cost of transfers is sufficiently low.

The rest of the paper is organized as follows. Section 2 presents a general framework explaining the non-cooperative and cooperative solutions to mitigate climate change. Section 3 shows the effect of asymmetries among countries on the outcomes of climate negotiations. Section 4 discusses the effects of pre-negotiations domestic policies and Section 5 concludes.

2 Model

2.1 Welfare Function

Consider the following welfare function of country $i \in N$:

$$W_i = B_i(Q) - C_i(q_i)\,, \tag{1}$$

where $N = \{1, 2, \ldots, n\}$ denotes a finite set of n countries. In the business-as-usual (BAU) scenario, each country has an initial unabated level of greenhouse gas (GHG) emissions e_i^o. If the country adopts a cooperative or non-cooperative climate policy to reduce emissions, its welfare function W_i can be defined as the difference between the benefits and costs of GHG abatement. Let q_i denote a particular abatement level of country i, and the vector of abatement levels of countries be $q = (q_1, q_2, ..., q_n)$. Because GHG emissions mix uniformly in the atmosphere, each country receives benefits not only from its own contribution, but also from other countries. Hence, the benefit function, which reflects the interdependence among countries, depends on global abatement Q, where $Q = \sum_{i \in N} q_i$. In contrast, the cost function depends only on the individual level of abatement $C_i(q_i)$.

The welfare function of each country is assumed to be strictly concave and twice differentiable. The benefit function increases in Q at a decreasing (or constant) rate, while the cost function increases in q_i at an increasing rate. Therefore, we have the following standard assumptions for all $i \in N$: $B_i^{'} > 0$, $B_i^{''} \leq 0$, $B_i(0) = 0$, and $C_i^{'} > 0$, $C_i^{''} > 0$ and $C_i(0) = 0$ (Finus, 2001; Hoel, 1991). These assumptions are sufficient for interior solutions. The abatement space is compact and convex $\mathbb{Q} = \mathbb{Q}_1 \times ... \times \mathbb{Q}_n$, where the abatement space of each country is $\mathbb{Q}_i \in [0, e_i^o]$ $\forall i \in N$.

2.2 Non-Cooperative Outcome

In the absence of cooperation, each country pursues its own self-interest and chooses an abatement level that maximises its individual welfare function (1) taking the abatement levels of other countries as given. The first-order condition of each country reads:

$$\begin{aligned} \frac{\partial W_i}{\partial q_i} &= B_i^{'}(Q) - C_i^{'}(q_i) = 0 \;\; \forall i \in N \\ \Leftrightarrow \;\; & B_i^{'}(Q) = C_i^{'}(q_i) \; . \end{aligned} \tag{2}$$

The above equation implies that the Nash equilibrium (NE) abatement level q_i^{*NE} of each country is determined when its marginal benefit is equal to its marginal cost. Obviously, the non-cooperative outcome is not Pareto-optimal because each country does not consider the benefits of other countries, which leads to the underprovision of GHG abatement levels required globally. Therefore, cooperation of countries is needed to internalise global benefits from abatement and to reach a Pareto-efficient outcome.

The welfare function of each country in the non-cooperative outcome is a function of the vector of the NE abatement levels, $W_i^{*NE}(q_1^{*NE},...,q_n^{*NE})$ $\forall i \in N$. Global abatement and welfare levels are given by $Q^{*NE} = \sum_{i\in N} q_i^{*NE}$ and $W^{*NE} = \sum_{i\in N} W_i^{*NE}$, respectively. The NE welfare level W_i^{*NE} is an important reference point in the negotiations among countries to decide whether joining a climate agreement is individually rational for each country. That is, a necessary condition for each country to accept cooperation is to obtain a welfare level under cooperation, W_i^{*C}, at least as high as its welfare level in the non-cooperative outcome.[3] Thus, individual rationality constraint requires: $W_i^{*C} \geq W_i^{*NE} \quad \forall i$.

2.3 Cooperative Outcomes (Pareto-optimal Outcomes)

All Pareto-optimal outcomes can be obtained from maximising the weighted sum of countries' welfare functions as follows:

$$\underset{q}{Max} \sum_{i\in N} \alpha_i W_i\left(q\right), \tag{3}$$

where the weight attached to each country is $\alpha_i \geq 0$ and $\sum_{i=1}^{n} \alpha_i = 1$. These weights can be interpreted as the power or bargaining position of countries in a global (social) welfare function (Escapa and Gutiérrez, 1997; Eyckmans and Cornillie, 2002; Eyckmans et al., 1993).

The first-order condition associated with problem (3) is given by:

$$\begin{aligned} &\sum_{j=1}^{N} \alpha_j B_j^{'}\left(Q\right) - \alpha_i C_i^{'}(q_i) = 0 \\ \Leftrightarrow &\sum_{j=1}^{N} \alpha_j B_j^{'}(Q) = \alpha_i C_i^{'}(q_i)\,. \end{aligned} \tag{4}$$

The above condition shows that the abatement level of each country under a cooperative outcome is such that the weighted marginal cost of each country is equal to the sum of weighted marginal benefits of all countries. This implies that each country internalises the benefits of other countries and the weighted marginal costs of countries are equal. Therefore, the ratio of weights of countries

[3] Individual rationality or profitability is only a necessary condition for countries to participate in an agreement. However, a sufficient condition requires that each country obtains at least what it would obtain if it remains outside the agreement while other countries continue to cooperate, which is referred to as the 'internal stability' condition due to d'Aspremont et al. (1983).

in global welfare function (3) is equal to the inverse ratio of their marginal costs of abatement,

$$\alpha_i C_i'(q_i) = \alpha_j C_j'(q_j) \ \ \forall i \neq j$$
$$\Leftrightarrow \ \frac{\alpha_i}{\alpha_j} = \frac{C_j'(q_j)}{C_i'(q_i)} \,. \tag{5}$$

Given that the welfare function of each country is strictly concave, the maximisation of (3) derives the whole set of Pareto-optimal outcomes for each vector of weights $\alpha = (\alpha_1, \alpha_2, ...\alpha_n)$. Consequently, we have:

- The slope of the line tangent to the Pareto frontier at any point is the ratio of the weights attached to countries in the global welfare function:$-\frac{\alpha_i}{\alpha_j}$.[4]
- For the welfare function (1), the absolute value of the slope of the line tangent to the Pareto frontier at any point is equal to the inverse ratio of marginal abatement costs of countries.

Different points on the Pareto frontier correspond to different cooperative solutions that could emerge from the negotiations. In the following two subsections, we consider three Pareto-optimal outcomes. We start with the first-best (FB) solution, where the socially (globally) optimal abatement level is determined by the unconstrained joint welfare maximisation. After this we consider two second-best (SB) solutions, where the abatement burdens are allocated among countries based on the constrained joint welfare maximisation or the Nash Bargaining Solution (NBS). In addition, we show how the relative weights of countries in (3) correspond to the three outcomes.

2.3.1 Unconstrained Joint Welfare Maximisation (FB Solution)

The individual abatement levels under the FB outcome follows from:

$$\underset{q_i}{Max} \sum_{j \in N} \left[B_j(Q) - C_j(q_j) \right], \tag{6}$$

which gives:

$$\sum_{j \in N} B_j'(Q) = C_i'(q_i)\,, \tag{7}$$

[4] See the proof in Engwerda (2005). However, for instructive purposes, I illustrate the proof for the welfare function (1) in Appendix A.1.

$$C_i'(q_i) = C_j'(q_j) \;\; \forall\, i \neq j\,. \tag{8}$$

Equation (7) shows that individual marginal costs of abatement are equal to the sum of marginal benefits in all countries, implying that global benefits from GHG abatement are fully internalised. Thus, in (8), the marginal abatement costs of countries are equalised implying that the abatement levels are allocated among countries cost-effectively.

The FB social optimum is attained if weights attached to countries in (3) are equal, i.e. $\alpha_i = \frac{1}{n}$. That is, the cost effectiveness condition in (8) implies that the absolute value of the slope of the line tangent to the Pareto frontier at the FB outcome is equal to one,

$$(\frac{\alpha_i}{\alpha_j})^{*Unc} = \frac{C_j'(q_j)}{C_i'(q_i)} = 1\,, \tag{9}$$

where the superscript *(Unc)* refers to the unconstrained maximisation.

Each country chooses its optimal abatement level q_i^{*Unc} from solving (7). Let the vector of abatement levels under the FB solution be $q^{*Unc} = (q_1^{*Unc}, q_2^{*Unc}, ..., q_n^{*Unc})$, and the global abatement level is $Q^{*Unc} = \sum_{i \in N} q_i^{*Unc}$. Inserting these equilibrium abatement levels into (1) gives individual and global welfare levels as W_i^{*Unc} and $W^{*Unc} = \sum_{i \in N} W_i^{*Unc}$, respectively.

A cost-effective allocation of abatement burdens as in (8) implies that countries with relatively flatter marginal cost functions should reduce emissions more than other countries. However, if those countries perceives benefits from global abatement to be sufficiently low, they would be worse off under cooperation than in the non-cooperative outcome and hence cooperation may not be individually rational for those countries.

One possible measure to compensate the losers from cooperation is to use side payments or transfers τ. Most of the literature assumes self-financed transfers where $\sum_{j \in N} \tau_j = 0$. The equilibrium individual welfare level can be modified with transfers to be $W_i^{Unc+\tau} = W_i^{*Unc} \pm \tau_i$.[5] We will analyse the effects of adding transfers in subsection 3.3. However, although transfers can greatly enhance cooperation among asymmetric countries, they are rarely observed in reality, in particular monetary transfers. Therefore, if transfers are not allowed, countries would depart from the FB outcome and choose different cooperative solutions.

[5]If a country receives transfers, then $\tau_i > 0$, while if it pays transfers, $\tau_i < 0$, so that $\sum_{j \in N} \tau_j = 0$.

2.3.2 Second-best Solutions

In this subsection, the allocation of abatement burdens among countries may follow from the constrained joint welfare maximisation or the Nash bargaining solution (NBS), which are referred to as second-best (SB) solutions.

1) Constrained Joint Welfare Maximisation

One possible outcome of the negotiations is to allocate abatement burdens such that no one country is made worse off under cooperation. Under this solution, countries maximise their joint welfare subject to a constraint that each country obtains at least as high as it would obtain in the non-cooperative outcome (Hoel, 1992). Therefore, the objective function can be written as:

$$\underset{q_i}{Max} \sum_{j \in N} [B_j(Q) - C_j(q_j)]$$

subject to

$$W_i = B_i(Q) - C_i(q_i) \geq W_i^{*NE} \ . \tag{10}$$

The corresponding Lagrangian is:

$$L = \sum_{j \in N} [B_j(Q) - C_j(q_j)] + \sum_{j \in N} \lambda_j [B_j(Q) - C_j(q_j) - W_j^{*NE}]. \tag{11}$$

The first-order condition can be summarised as follows (see Appendix A.2):

$$\sum_{j \in N} B_j^{'}(Q) + \frac{\sum_{j \in N} \lambda_j B_j^{'}(Q) - \lambda_i \sum_{j \in N} B_j^{'}(Q)}{1 + \lambda_i} = C_i^{'}(q_i)\,. \tag{12}$$

Simplifying the above condition, we can obtain the following:

$$(1 + \lambda_i) C_i^{'}(q_i) = (1 + \lambda_j) C_j^{'}(q_j),\ \forall\ i \neq j\,. \tag{13}$$

$$\frac{1 + \lambda_i}{1 + \lambda_j} = \frac{C_j^{'}(q_j)}{C_i^{'}(q_i)}\,. \tag{14}$$

Compared to (8), the cost-effectiveness condition clearly breaks down in the above equation since marginal costs are not equalised. Thus, the inverse ratio of marginal abatement costs of countries depends on the ratio of the Lagrangian multipliers associated with their individual rationality constraint.

Since the outcome of the above solution lies on the Pareto frontier, there is a vector of weights in (3) corresponding to this outcome. These weights can be written in terms of each country's Lagrangian multiplier. If we compare the first-order condition in (12) with (4), and compare (14) with (5), the ratio of countries' weights can be written in terms of the Lagrange multiplier as follows:

$$(\frac{\alpha_i}{\alpha_j})^{*Con} = \frac{1+\lambda_i}{1+\lambda_j}\,, \tag{15}$$

where the superscript *(Con)* refers to the constrained maximisation.

Therefore, those countries whose Lagrangian multiplier is relatively larger should be attached a higher weight relative to other countries in the global welfare function and have lower marginal costs in equilibrium (Eyckmans, 2009).

Let the vector of the abatement levels under the constrained maximisation solution be $q^{*Con} = (q_1^{*Con}, q_2^{*Con}, ..., q_n^{*Con})$, and the global abatement level is $Q^{*Con} = \sum_{i \in N} q_i^{*Con}$. Inserting these equilibrium abatement levels into (1) gives individual and global welfare levels as W_i^{*Con} and $W^{*Con} = \sum_{i \in N} W_i^{*Con}$, respectively. For those countries whose constraint is binding, they would obtain the same welfare level as under the non-cooperative outcome and hence all gains are captured by other countries which have a non-binding constraint.

2) Nash Bargaining Solution (NBS)

The outcome of climate negotiations could also follow from the Nash bargaining solution which maximises the product of the gains from cooperation as follows:

$$\underset{q_i}{Max} \prod_{i \in N} \left(B_i(Q) - C_i(q_i) - W_i^{*NE}\right), \tag{16}$$

where W_i^{*NE} represents the 'threat or disagreement point' as defined in subsection 2.2. That is, if countries fail to negotiate an agreement, they end up in the non-cooperative NE outcome (Hoel, 1991).

Equation (16) corresponds to:

$$\underset{q_i}{Max} \sum_{j \in N} \log \left(B_j(Q) - C_j(q_j) - W_j^{*NE}\right). \tag{17}$$

The first-order condition can be written as follows (see details in Appendix A.3):

$$G_i \sum_{j \in N} \frac{1}{G_j} B_j^{'}(Q) = C_i^{'}(q_i)\,, \tag{18}$$

where $G_i = W_i - W_i^{*NE}$ are the gains from cooperation.

The above condition states that the ratio of each country's marginal cost to its gains from cooperation is equal to the sum of the ratio of marginal benefits of countries to their gains from cooperation. Therefore, we can write the following:

$$\frac{C_i^{'}(q_i)}{G_i} = \frac{C_j^{'}(q_j)}{G_j} \quad ,\forall i \neq j\,. \tag{19}$$

The allocation of the abatement levels under the NBS is cost-ineffective because marginal costs of countries are not equalised. Countries' relative marginal costs are equal to the ratio of their gains from cooperation in equilibrium:

$$\frac{G_i}{G_j} = \frac{C_i^{'}(q_i)}{C_j^{'}(q_j)} \,,\forall i \neq j\,. \tag{20}$$

One of the main axiomatic properties of the NBS is Pareto-optimality (Nash Jr, 1950). Thus, the vector of weights in (3) corresponding to the NBS outcome is given by:

$$(\frac{\alpha_i}{\alpha_j})^{*NBS} = \frac{G_j}{G_i} \quad ,\forall i \neq j\,, \tag{21}$$

where the superscript *(NBS)* refers to the Nash bargaining solution.

From (5) and (20), we have the ratio of weights of countries corresponding to the NBS outcome equal to the inverse ratio of their gains from cooperation or, in other words, the weighted gains from cooperation are equal (Douven and Engwerda, 1995; Engwerda, 2005).

Let the vector of the abatement levels under the NBS be $q^{*NBS} = (q_1^{*NBS}, q_2^{*NBS}, ..., q_n^{*NBS})$, and the global abatement level is $Q^{*NBS} = \sum_{i \in N} q_i^{*NBS}$. Inserting these equilibrium abatement levels into (1) gives individual and global welfare levels as W_i^{*NBS} and $W^{*NBS} = \sum_{i \in N} W_i^{*NBS}$, respectively.

By comparing the first-order conditions given in (7), (12) and (18), it is clear that the three solutions yield the same outcome if countries are symmetric. In such cases, the individual rationality constraint of all countries is non-binding, i.e. $\lambda_i = 0\ \forall i \in N$, and equation (12) converges to (7). Similarly, if countries are symmetric, they have the same gain from cooperation under the NBS $G_i = G_j = ... = G_n$. Therefore, (18) also converges to (7). Therefore, comparing the three outcomes is interesting only if countries have asymmetric benefit and cost functions as will be shown in the next section.

3 Asymmetry and Climate Negotiations Outcomes

For the rest of the paper, we will consider negotiations between two asymmetric countries or regions, $i = 1, 2$, choosing their abatement levels cooperatively or non-cooperatively. This helps us to compare the three Pareto-optimal outcomes under different types of asymmetry.

In the absence of transfers, we need to focus our attention on a setting where certain countries become worse off under the FB outcome than under no cooperation due to asymmetries. For this, we assume, without loss of generality, that the FB outcome is not individually rational for country 2.

Assumption 1(a):

If transfers are not allowed, country 2 is worse off under cooperation using joint welfare unconstrained maximisation, i.e. the individual rationality constraint of country 2 is binding, $\lambda_2 > \lambda_1 = 0$.

Given the above assumption, we can illustrate the trade-off between individual rationality and efficiency in negotiations on the Pareto frontier as shown in Figure 1. The FB solution is represented by point F on the Pareto frontier. Given Assumption 1(a), the FB outcome cannot be reached. The range of the possible outcomes for a Pareto-improving agreement lies between two points: C_1 and C_2.

Point C_2 corresponds to the outcome under the constrained joint welfare maximisation if the individual rationality constraint of country 2 is binding.[6] This point represents the maximum (minimum) weight country 1 (2) would have relative to country 2 (1) for both of them to cooperate. At this point, the slope of the line tangent to the Pareto frontier becomes flatter than under the unconstrained outcome: $(\frac{\alpha_1}{\alpha_2})^{*Con} < (\frac{\alpha_1}{\alpha_2})^{*Unc} = 1$. Point C_2 is associated with the minimum global welfare loss that can be incurred if we depart from the FB outcome for a Pareto-improving agreement. At this outcome, all gains from cooperation go to country 1, while country 2 obtains the same welfare level as under no cooperation.

[6]Point C_1 corresponds to the outcome under the constrained joint welfare maximisation if the constraint of country 1 is binding.

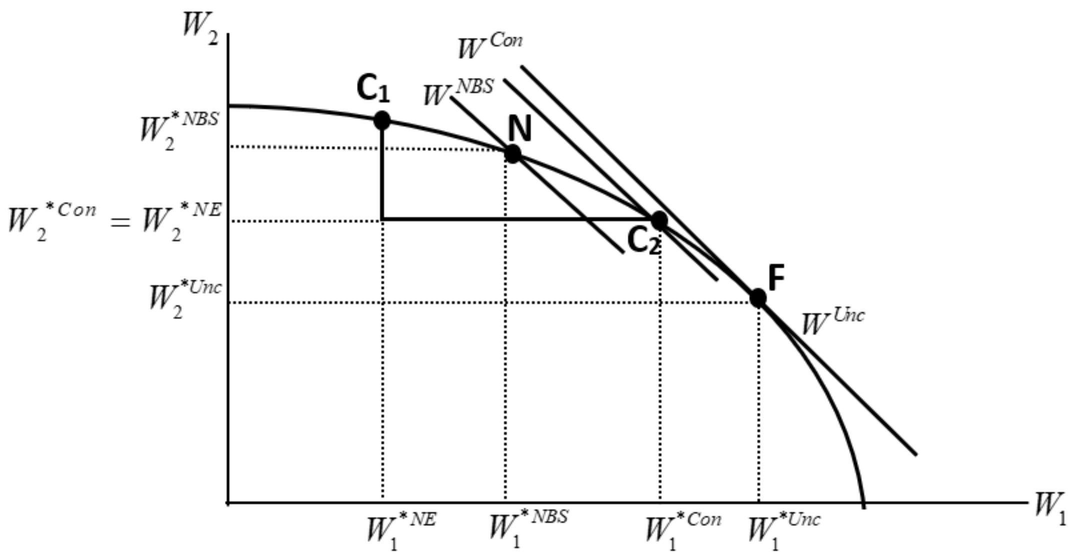


Figure 1: Individual Rationality versus Efficiency in Climate Change Negotiations

Point N corresponds to the NBS, where both countries gain from cooperation.[7] The slope of the line tangent to the Pareto frontier at the NBS outcome is even flatter such that we have $(\frac{\alpha_1}{\alpha_2})^{*Unc} > (\frac{\alpha_1}{\alpha_2})^{*Con} > (\frac{\alpha_1}{\alpha_2})^{*NBS}$. This clearly implies that, across the three solutions, country 1 obtains the highest welfare level under the unconstrained maximisation, followed by the constrained maximisation and then the NBS, $W_1^{*Unc} > W_1^{*Con} > W_1^{*NBS}$, while the opposite is true for country 2, $W_2^{*Unc} < W_2^{*Con} < W_2^{*NBS}$. In addition, compared to the FB outcome, the NBS is associated with a larger global welfare loss if the alternative is the constrained maximisation.

As shown in (8), marginal costs of abatement are equalised across countries under the FB outcome. However, given Assumption 1(a), country 2 will not accept cooperation. In such cases, country 1 has higher marginal costs than country 2 under the SB solutions.

Lemma 1. *Given Assumption 1(a), i. country 1 has higher marginal costs than country 2 under the constrained maximisation and the NBS,* $C_1'(q_1^{*Con}) > C_2'(q_2^{*Con})$ *and* $C_1'(q_1^{*NBS}) > C_2'(q_2^{*NBS})$.

ii. Country 1 has larger gains than country 2 under the NBS, $G_1 > G_2$.

Proof. Follows from (14), (15) and (20), (21), given Assumption 1(a) that $\lambda_2 > \lambda_1 = 0$ and $(\frac{\alpha_1}{\alpha_2})^{*NBS} = \frac{G_2}{G_1} < 1$. □

[7]This point is determined by the tangency between the Pareto frontier and a rectangular hyperbola with the origin the disagreement point, which is the NE welfare levels in Figure 1. For the geometric proof of the NBS, see Binmore (1998), pp.78-80 and Friedman (1986), pp.157-158.

3.1 Specific Welfare Function

In order to determine the conditions under which the FB solution is not mutually profitable without transfers, we use a specific welfare function. Following many papers in the IEA literature (e.g., Finus and McGinty, 2019; Hoel, 1992; Pavlova and De Zeeuw, 2013), we use a linear benefit-quadratic cost welfare function:

$$W_1 = bQ - \frac{c}{2}(q_1)^2, \tag{22}$$

$$W_2 = \beta bQ - \frac{\gamma c}{2}(q_2)^2, \tag{23}$$

where one unit of GHG abatement q_i generates to country 1 and country 2 constant marginal benefits of b and βb , respectively; and marginal costs of cq_1 and γcq_2, respectively. Assumptions regarding the values of β and γ will be stated below according to different types of asymmetry, however, we assume all parameters are positive, $b, c, \beta, \gamma > 0$.

Different marginal benefits among countries can be interpreted in terms of their vulnerability to climate change or their willingness to pay for emission reductions.[8] Whereas different marginal cost functions might reflect differences across countries with respect to their technological efficiency or their fossil fuel consumption. As shown empirically, using the Stability of Coalition (STACO) model (e.g., Altamirano-Cabrera et al., 2008; Finus, 2008), large industrialised countries similar to those in the European Union, the Unites States (USA) and Japan are the main beneficiaries of global abatement, whereas energy exporting and Asian countries and Brazil receive the lowest benefits.[9] On the abatement cost side, both China and USA have the flattest marginal abatement cost curves, whereas Brazil and Japan have the steepest curves.

Using the welfare functions in (22) and (23), the equilibrium abatement and welfare levels of countries under the non-cooperative and the cooperative outcomes are presented in Appendix B.1.

In what follows, we show the effect of the asymmetry among countries in terms of their evaluation of benefits and costs of GHG abatement on the negotiations to reach a climate agreement. Taking symmetry as a benchmark, $\beta = 1$ and

[8] It can reflect the differences in income levels among countries. High-income countries are willing and able to pay more for improved environmental quality than low-income countries (Copeland and Taylor, 2005).

[9] STACO (Stability of Coalitions) model investigates the formation and stability of international environmental agreements. This model captures twelve world regions and considers a period of 100 years, see Dellink et al. (2004).

$\gamma = 1$, both countries will gain from cooperation under the FB solution and hence a cost-effective allocation of abatement burdens is achievable. However, asymmetries could change the incentives for countries to cooperate. Proposition 1 shows that a negative correlation between the cost and benefit sides is more likely to lead to the FB outcome.

Proposition 1. The First-best Outcome versus No Cooperation

i. The two countries have incentives to cooperate under the FB solution without transfers if $\frac{1}{2} < \beta^2\gamma < 2$ or, to put it differently, if a) $\beta^2 < \frac{2}{\gamma}$ or b)$\beta^2 > \frac{1}{2\gamma}$.

ii. Let the potential gains from cooperation under the FB outcome $\Delta W = W^{*Unc} - W^{*NE}$, then $\Delta W = \frac{b^2}{2c}\frac{(\beta^2\gamma+1)}{\gamma}$.

Proof. See Appendix B.1. □

In the absence of transfers, the FB social optimum is individually rational if the asymmetry on the cost side is compensated by a large inverse asymmetry on the benefit side. That is, if country 2 has a flat marginal cost function, it would accept a cost-effective allocation of abatement burdens (i.e. under the FB outcome) only if its marginal benefits from abatement are sufficiently large. This type of asymmetry is referred to in the literature as 'negative covariance', i.e. countries with higher marginal benefits of abatement also have flatter marginal cost functions (Finus and McGinty, 2019; McGinty, 2007). If transfers are not allowed, Finus and McGinty (2019) explain that the compensatory role of transfers could work through this type of asymmetry. However, if we look at the negotiations between developed and developing countries on climate change, we can notice the opposite type of asymmetry. For instance, and as mentioned above, most of the developed countries perceive benefits from abatement more importantly than developing countries, but face steeper marginal cost curves which is refereed to as a 'positive covariance' between benefits and costs.

Note that the global welfare gains that could be achieved from cooperation under the FB outcome, where global welfare is maximised, increase in the benefit parameters, b and β, while decrease in the cost parameters c and γ.

In Proposition 1, violation of condition a), i.e. $\beta > \sqrt{\frac{2}{\gamma}}$, or equivalently $\gamma > \frac{2}{\beta^2}$, implies that the constraint of country 1 is binding $\lambda_1 > 0$ and $\lambda_2 = 0$. In contrast, violation of condition b), i.e. $\beta < \sqrt{\frac{1}{2\gamma}}$, or equivalently $\gamma < \frac{1}{2\beta^2}$, implies that the constraint of country 2 is binding $\lambda_1 = 0$ and $\lambda_2 > 0$. Following Assumption 1(a), we impose the following constraint on the benefit and cost parameter values.

Assumption 1(b):

Let $\beta < \sqrt{\frac{1}{2\gamma}}$, or equivalently, $\gamma < \frac{1}{2\beta^2}$. That is, country 2 has no incentive to cooperate under the FB solution if transfers are not allowed.

In the previous section, we showed that for each Pareto-optimal outcome, there are corresponding relative weights associated to countries in the global welfare function. Using (22) and (23), we can show that these weights depend on the gains from cooperation under the FB outcome, which in turns depend only on β and γ. That is, the relative weights corresponding to the constrained maximisation and the NBS are $(\frac{\alpha_1}{\alpha_2})^{*Con} = 2^{1/3}\left(\beta^2\gamma\right)^{1/3}$ and $(\frac{\alpha_1}{\alpha_2})^{*NBS} = \left(\beta^2\gamma\right)^{1/3}$, respectively. In Proposition 1, the lower the value of $\beta^2\gamma$, the larger the gains of country 1 from cooperation under the FB solution and the lower is its relative weight under the SB Solutions. For six of the world regions, Escapa and Gutiérrez (1997) show, using numerical simulations, that the relative power weights of countries in negotiations are negatively related to their net benefits under the FB solution.

3.2 Comparison of the Equilibrium Outcomes

In this subsection, we analyse how the type of asymmetry affects the distribution of the abatement burden between countries under the negotiated outcomes. In addition, we evaluate the impacts of asymmetries on global abatement and global welfare. Although we do the comparison between the three cooperative outcomes, recall that, given Assumption 1, the FB outcome cannot be achieved.

Before turning to the allocation of burdens among countries, the following result shows that the abatement efforts of country 1 increase as we move from the FB to the constrained outcome, and increase even further under the NBS. That is, as the relative weight of country 1 decreases, its contribution to reduce emissions increases.[10] On the other hand, country 2, for which the constraint is binding, abates less under the two SB solutions compared to the FB solution.

Corollary 1. Ranking of Equilibrium Abatement of Countries

i. $q_1^{*NBS} > q_1^{*Con} > q_1^{*Unc} \Longrightarrow C_1'\left(q_1^{*NBS}\right) > C_1'\left(q_1^{*Con}\right) > C_1'\left(q_1^{*Unc}\right)$.

ii. $q_2^{*Unc} > q_2^{*Con} > q_2^{*NBS} \Longrightarrow C_2'\left(q_2^{*NBS}\right) < C_2'\left(q_2^{*Con}\right) < C_2'\left(q_2^{*Unc}\right)$.

Proof. See Appendix B.2. □

[10]As will be mentioned in Section 4, Hoel (1991) shows that as the position of a country is weakened in negotiations, it incurs more abatement efforts.

In what follows, we consider the impacts of different types of asymmetry. Most of the literature considers asymmetry either on the benefit or the cost side, hence, we start our analysis with one-sided asymmetry. On the one hand, if the two countries are asymmetric on the benefit side only, i.e. $\gamma = 1$, Assumption 1(b) implies that $\beta < \sqrt{\frac{1}{2}} \simeq 0.707$. Thus, country 1 has a higher marginal benefit than country 2. Symmetric cost functions implies that both countries contribute the same level of abatement under the FB outcome. Whereas in the two SB solutions, the allocation of burdens among countries depends on their marginal benefits of abatement, where the country with higher marginal benefits abates more.[11] On the other hand, if the two countries are asymmetric on the cost side only, i.e. $\beta = 1$, country 2 has a flatter marginal cost curve, $\gamma < \frac{1}{2}$ (see Assumption 1(b)). A cost-effective allocation of abatement implies that the country which has a relatively flatter marginal cost function always contributes more under the FB solution. This is likewise the allocation of the abatement efforts under the SB solutions. However, although country 2 contributes more than country 1 under the three outcomes, we show in Corollary 1 that it contributes the highest level of abatement under the FB outcome. Thus, country 2 shifts some of its abatement burden to country 1 under the SB solutions.

Corollary 2. Allocation of Abatement Burdens: One-sided Asymmetry

i. If both countries have symmetric cost functions, $\gamma = 1$, they contribute the same level of abatement under the FB solution: $q_1^{*Unc} = q_2^{*Unc}$, while country 1 abates more than country 2 under the SB solutions: $q_1^{*Con} \left(q_1^{*NBS}\right) > q_2^{*Con} \left(q_2^{*NBS}\right)$.

ii. If both countries have symmetric benefit functions, $\beta = 1$, country 2 abates more than country 1 under the three solutions.

Proof. See Appendix B.3. □

More realistically, the following result considers two-sided asymmetry, i.e. countries differ on both the benefit and cost sides. As mentioned before, the FB solution implies that countries with a relatively flatter marginal cost function contribute more than other countries, irrespective of the differences in their marginal benefits of abatement. On the contrary, the relation between asymmetries on both sides determines the allocation of abatement burdens among countries under the SB solutions, which leads to a more symmetric distribution of the gains from cooperation.

[11] In an emission game, and if countries are asymmetric on the damage (benefit in the abatement game) side, see for instance Hoel (1992) under the constrained joint welfare maximisation, and Boom (2006) under the NBS.

Corollary 3. Allocation of Abatement Burdens: Two-sided Asymmetry

- *If there is a negative $\beta - \gamma$ covariance, we may have:*

  i. $\beta < 1$ *and* $\gamma > 1$*, where country 1 contributes more than country 2 under the three outcomes.*

  ii. $\beta > 1$ *and* $\gamma < \frac{1}{2}$*, where country 2 contributes more than country 1 under the three outcomes.*

- *If there is a positive $\beta - \gamma$ covariance, we have $\beta < 1$ and $\gamma < 1$, where,*

  i. $q_1^{*Unc} < q_2^{*Unc}$.

  ii. $q_1^{*Con} \geq (<)\, q_2^{*Con}$ *if* $\beta \leq (>)\, \frac{1}{\sqrt{2}}\gamma$.

  iii. $q_1^{*NBS} < q_2^{*NBS}$ *for all* $\gamma > \overline{\gamma} \simeq 0.79$ *and* $q_1^{*NBS} \geq (<)\, q_2^{*NBS}$ *if* $\beta \leq (>)\, \gamma$ *for all* $\gamma < \overline{\gamma}$.

Proof. See Appendix B.4. □

We analyse two types of asymmetry. First, a negative covariance between benefit and cost parameters (negative $\beta - \gamma$ covariance), where the country with a higher marginal benefit has a flatter marginal cost curve and vice-versa, i.e. the country with a lower marginal benefit has a steeper marginal cost curve. On the one hand, if we assume $\gamma > 1$, Assumption 1(b) implies that $\beta < 1$. Thus, country 1 has a higher marginal benefit and a flatter marginal cost curve than country 2 and contributes more than country 2 under the three outcomes. Nevertheless, across the three outcomes, country 1's abatement level would be the highest under the NBS (see Corollary 1). On the other hand, if we assume $\gamma < 1$, then $\beta > 1$ can only hold if $\gamma < \frac{1}{2}$ given Assumption 1(b). This implies that country 2 has a higher marginal benefit and a flatter marginal cost curve and contributes more than country 1.

Second, a positive covariance between benefits and costs (positive $\beta - \gamma$ covariance), where a country with a lower marginal benefit has a flatter marginal cost curve, and vice-versa, i.e. a country with a higher marginal benefit has a steeper marginal cost curve. Given Assumption 1(b), we cannot have both $\gamma > 1$ and $\beta > 1$ at the same time. That is, a country with higher marginal benefits and steeper marginal costs has a non-binding constraint (like country 1). For this reason, we consider only the case where, $\beta < 1$ and $\gamma < 1$. That is, country 2 has a flatter marginal cost curve and lower marginal benefits than country 1.[12]

[12]Note that for $\frac{1}{2} < \gamma < 1$, Assumption 1(b) is sufficient to guarantee that $\beta < 1$.

These conditions imply that country 2 always contributes more than country 1 under the FB solution. However, under a more symmetric burden sharing rules, country 2 may contribute more or less than or equal to country 1, depending on the benefit and cost asymmetries.

Corollary 1 has shown that moving from the FB solution to either the constrained outcome or the NBS, country 1 (2) contributes more (less) to mitigate emissions. Therefore, the effect on global abatement is a priori ambiguous.

Corollary 4. Ranking of Equilibrium Global Abatement

- *For* $\gamma = 1$*:*

  i. $Q^{*Unc} > Q^{*Con}$ *and* $Q^{*Unc} > Q^{*NBS}$*;*

  ii. $Q^{*Con} > (<)\, Q^{*NBS}\ \forall \beta > (<)\, 0.5$*.*

- *For* $\beta = 1$*:* $Q^{*Unc} > Q^{*Con} > Q^{*NBS}$ *.*

- *For a negative* $\beta - \gamma$ *covariance:*

  a) $\beta < 1$ *and* $\gamma > 1$*: i.* $Q^{*Unc} > Q^{*Con}\ \forall \gamma < 2$*, while* $Q^{*Unc} \gtrless Q^{*Con}$ *if* $\gamma > 2$*;*

  ii. $Q^{*Unc} > Q^{*NBS}\ \forall \gamma < 1.26$*, while* $Q^{*Unc} \gtrless Q^{*NBS}$ *if* $\gamma > 1.26$*;*

  iii. $Q^{*Con} > (<)\, Q^{*NBS}$ *if only if* $\beta > (<)\, 0.5\gamma\ \forall \gamma < 1.26$*, while* $Q^{*Con} < Q^{*NBS}\ \forall \gamma > 1.26$*.*

  b) $\beta > 1$ *and* $\gamma < 0.5$*:* $Q^{*Unc} > Q^{*Con} > Q^{*NBS}$*.*

- *For a positive* $\beta - \gamma$ *covariance,* $\beta < 1$ *and* $\gamma < 1$*:*

  i. $Q^{*Unc} > Q^{*Con}$ *and* $Q^{*Unc} > Q^{*NBS}$*;*

  ii. $Q^{*Con} \geq (<)\, Q^{*NBS}\ \forall \beta \geq (<)\, 0.5\gamma$*.*

Proof. See Appendix B.5. □

Corollary 4 shows that total abatement generally decrease as we depart from the FB solution to the SB solutions for most types of asymmetry. An exception is the case of a negative $\beta - \gamma$ covariance under which country 1 has a higher marginal benefit and a flatter marginal cost curve. In this case, if the asymmetry among countries is large on both sides, we may have a higher global abatement level under the SB solutions than under the FB outcome. With respect to the two SB solutions, comparing the global abatement level is less straightforward since it depends on the degree of asymmetry.

From the above results, we can conclude that if the asymmetry among countries is sufficiently large, the NBS will lead not only to a more symmetric distribution of the gains from cooperation, but also to higher global abatement compared to the constrained outcome. However, this goes along with a higher global welfare loss. In addition, only if the country for which the individual rationality constraint is not binding has a higher marginal benefit and a flatter marginal cost curve, global abatement could be higher under the SB solutions compared to the FB solution. Though higher global abatement is also associated with a global welfare loss.

The degree and type of asymmetry among countries also affect the global welfare loss as we move away from the FB solution. As shown in Figure 1, the parallel lines W^{*Unc}, W^{*Con} and W^{*NBS} represent the global welfare level at each outcome, where it decreases as we depart from the FB solution moving up along the Pareto curve. In the following analysis, we evaluate the two SB solutions in terms of global welfare using a relative measure called the closing the gap index (CGI) as suggested by Eyckmans and Finus (2006). In our context, the CGI measures to which extent the SB solutions close the gap between the FB solution and the non-cooperative outcome:

$$CGI^{Con} = \frac{W^{*Con} - W^{*NE}}{\Delta W} \cdot 100 \,\&\, CGI^{NBS} = \frac{W^{*NBS} - W^{*NE}}{\Delta W} \cdot 100\,, \qquad (24)$$

with ΔW as defined in Proposition 1.

The CGI in (24) depends only on β and γ.[13] We find that for all types of asymmetry mentioned above, the CGI decreases as the degree of asymmetry among countries increases. As usual, we start our analysis with one-sided asymmetry. On the benefit side, i.e. $\gamma = 1$ and $\beta < \sqrt{1/2}$, if the β-asymmetry increases, (i.e. β decreases), the CGI decreases, $\frac{\partial CGI}{\partial \beta} > 0$. However, this goes along with a decrease in the potential gains achieved from the FB outcome, $\frac{\partial \Delta W}{\partial \beta} > 0$. In contrast, on the cost side, i.e. $\beta = 1$ and $\gamma < \frac{1}{2}$, if the γ-asymmetry increases (i.e. γ decreases), the CGI also decreases, $\frac{\partial CGI}{\partial \gamma} > 0$, however the global welfare gap increases, $\frac{\partial \Delta W}{\partial \gamma} < 0$.

Consider now our two cases of the negative $\beta - \gamma$ covariance. On the one hand, if $\beta < 1$ and $\gamma > 1$, increasing the asymmetry can be obtained by a marginal decrease in β combined with a marginal increase in γ, i.e. $\beta - \epsilon\,(\gamma + \epsilon)$. In such cases, both the CGI and the ΔW decrease in the degree of asymmetry. Recall from above that ΔW decreases as β decreases and as γ increases. On the

[13]See Appendix B.6 for details.

other hand, if $\beta > 1$ and $\gamma < 1$, increasing negative $\beta - \gamma$ covariance can be obtained by a marginal increase in β combined with a marginal decrease in γ, i.e. $\beta + \epsilon\,(\gamma - \epsilon)$. As a result, the ΔW increases in the degree of asymmetry though the CGI decreases. For the positive $\beta - \gamma$ covariance case, i.e. $\beta < 1$ and $\gamma < 1$, asymmetry is increased by a marginal decrease in both β and γ, i.e. $\beta - \epsilon\,(\gamma - \epsilon)$. The effect of this type of asymmetry is less straightforward on the potential gains from cooperation. Starting from $\gamma = 1$, and $\beta = \sqrt{1/2}$ (see Assumption 1(b)), we find that ΔW first decreases and then increases in the degree of asymmetry. More precisely, when γ sufficiently decreases, its positive impact on ΔW outweighs the negative impact of low β. In any case, the CGI decreases in the degree of asymmetry.

3.3 The Role of Transfers in Negotiations

We have assumed so far that transfers are not allowed, and hence countries negotiate only about the allocation of the abatement burdens. However, with transfers, negotiations deal over both side payments and abatement levels. Therefore, in this subsection, we compare the incentives of countries if the FB solution can be made mutually profitable with the availability of transfers. It is common in the literature to assume that countries distribute the gains from cooperation equally (e.g., Botteon and Carraro, 1997; Carraro et al., 2006). The welfare level of each country with transfers can be written as:[14]

$$W_i^{*Unc+\tau} = W_i^{*NE} + \omega_i(\mathit{\Delta} W)\ \forall i = 1, 2\,, \tag{25}$$

with $\omega_i = \frac{1}{2}$, and $\mathit{\Delta} W$ as defined in Proposition 1.

Assumption 2:

If countries cooperate under the FB solution, they share the total gains from cooperation equally.

Taken together, Assumption 1 and 2 imply that country 1 gives side payments to compensate country 2. Hence, in what follows, we focus more on whether and under which conditions country 1 will prefer to pay transfers. Given the ranking of the equilibrium abatement levels in Corollary 1, country 1 chooses either to incur

[14]The transfer scheme we use in (25) is the NBS, under which each country obtains its non-cooperative NE welfare level and receives an equal share of the total gains from cooperation. Other transfers schemes are the Shapely Value and the Chander and Tulken transfer schemes. See footnote 1.

larger abatement burden under the SB than the FB outcome or to pay transfers to achieve the FB outcome. Both options imply a redistribution of the gains from country 1 to country 2 either in terms of reallocation of abatement burdens or direct transfers. In order to gain insights into the preferences of countries after allowing transfers, we consider only one-sided asymmetry.

Corollary 5. *Given Assumption 1, and if countries are asymmetric either on the benefit or the cost side, country 1 is more likely to give transfers under the FB outcome than to choose a SB outcome (either the constrained joint welfare maximisation or the NBS) only if the asymmetry among countries are sufficiently large.*

Proof. See Appendix B.7. □

On the one hand, if countries are asymmetric on the benefit side only, i.e.$\gamma = 1$ and $\beta < 0.707$ (recall Assumption 1(b)), we find that as long as the β-asymmetry is not very large, country 1 is better off under the SB outcome than giving transfers under the FB outcome. More specifically, the value of β which makes country 1 indifferent between the two outcomes is $\beta \simeq 0.49$ for the NBS and $\beta \simeq 0.14$ for the constrained maximisation outcome. However, if β becomes sufficiently small, implying larger benefit-asymmetry, country 1 prefers to choose the FB outcome with transfers. With respect to country 2, it is always better off under the FB with transfers. On the other hand, if countries are asymmetric on the cost side only, i.e. $\beta = 1$ and $\gamma < 0.5$, country 1 will also choose the SB solutions as long as the asymmetry among countries is not very large. We find that country 1 is indifferent between the two outcomes if $\gamma \simeq 0.24$ under the NBS and $\gamma \simeq 0.019$ under the constrained outcome. It is straightforward to show that even if we relax Assumption 2 and allow for unequal distribution of the gains from cooperation, country 1 is more likely to choose the FB outcome with transfers if the asymmetry among countries is large. That is, there is a positive relation between β or γ and ω_1. In other words, the larger the degree of asymmetry, the more likely that country 1 would give transfers, even for a smaller share of the surplus than country 2.

4 Pre-Negotiation Domestic Policies

We have shown in the previous section that climate negotiations outcomes depend greatly on the asymmetry among countries and whether additional instruments such as side payments are available. However, domestic policies such as unilateral abatement actions, investment and adaptation could also affect the disagreement welfare level of nations (i.e., non-cooperative status quo) which in turn affect the negotiations.

The impact of a unilateral emission reduction action on the outcomes of negotiations was analysed by Hoel (1991). He shows that if a country (say country 1 in our model) undertakes a unilateral emission reduction action in excess of its NE output level (q_1^{*NE}), the non-cooperative outcomes (the disagreement point) shifts in favour of country 2. Consequently, the new outcome of the NBS implies a reduction in the relative weight (bargaining position) of country 1 (for illustration, it can be any point from C_1 and N in Figure 1). As a result, the cooperative welfare level of country 1 (2) decreases (increases) and country 1 (2) incurs higher (lower) abatement efforts than an initial situation without commitment to a unilateral action. He finds that the change in global emissions depends on the ratio between the abatement cost parameters C_1''/C_2''. More precisely, in our notations, if $\frac{1}{\gamma}$ is high (low) global emissions will increase (decrease) as a result of a unilateral action.

Anticipating climate change negotiations could affect the countries' domestic policies. From this perspective, Buchholz and Konrad (1994), assuming symmetric countries, show that countries have a strategic incentive to choose technologies which reduce emissions at high per unit cost. They explain that having a relatively higher cost of abatement allows the country to commit to a lower abatement level either under no cooperation or under a cooperative agreement. If countries are symmetric on the benefit side, it is clear from (A.11) that the country with a steeper marginal cost curve (country 1 if $\gamma < 1$) contributes less and gains more under no cooperation than the other country. In addition, the country which chooses higher abatement costs would shift the disagreement point and the NBS in its favour. This leads to what is known in the literature by the 'hold-up' problem which may result in underinvestment of clean technology (Harstad, 2016). Adaptation to climate change could also affect the negotiations on mitigation. Zehaie (2009) shows that undertaking adaptation actions before negotiating the abatement levels could give countries a strategic advantage (stronger position) in bargaining such that they would secure lower abatement burden. The main intuition is that adaptation reduces the country's vulnerability to climate change (i.e. its benefits from abatement decrease), and hence its future abatement efforts. If countries are symmetric on the cost side, it is also clear from (A.11) that if $\beta < 1$, country 2 contributes less under no cooperation, but also under the SB agreements (see Corollary 2). Therefore, countries can shift their abatement burdens to others through adaptation.

Trade policies such as border carbon adjustments (BCAs) could also affect climate change negotiations (Helm et al., 2012). BCAs adjust the differences between national carbon prices through import tariffs or export rebates. These policies have been recently proposed to support unilateral actions of environmentally friendly countries or as trade sanctions to enforce climate agreements. If we go back to Section 3, assume that country 1, which gains from cooperation under the FB outcome (recall Assumption 1(a)), represents a developed or an environmentally more concerned country. Let point N in Figure 1 be the initial cooperative outcome given that BCAs are not available at the disagreement point. If country 1 chooses to support its non-cooperative welfare level by imposing BCAs against country 2, BCAs may shift the disagreement point and the cooperative outcome in favour of country 1. That is, the outcome of the negotiations could be any point between N and F in Figure 1 implying stronger bargaining position for country 1 and a higher global welfare level. Of course, at this abstract level, we cannot determine exactly the effects of BCAs on the non-cooperative outcome and under which conditions the FB solution (at point F) could be attained as a cooperative outcome. In order to analyse the detailed impacts of these policies, trade between countries should be modelled explicitly.

5 Conclusions

Solving the climate change problem requires cooperation among countries to internalise global benefits from GHG abatement. A Pareto-efficient solution for climate change can be reached through maximising the weighted sum of the welfare functions of all countries. The relative weights attached to countries' welfare functions lead to different cooperative outcomes with different allocation of abatement burdens. The incentives for countries to cooperate depend on their benefits and costs associated with abatement. However, benefits and costs of abatement differ among countries to such extent that some countries could be worse off under a cooperative outcome than under no cooperation implying that cooperation might not prove individually rational for them.

The first-best (FB) outcome can be achieved if the welfare weights attached to countries are equal, implying a cost-effective allocation of abatement burdens among countries. However, if transfers are not allowed to redistribute the gains achieved from cooperation, the FB solution may not be reached due to large asymmetries among countries. In such cases, countries would negotiate second-best (SB) agreements. Under the SB solutions, a lower relative weight should be attached to countries which gain more from cooperation. Despite the fact that the distribution of the potential gains from cooperation is more symmetric under the SB than the FB outcome, this is associated with a global welfare loss. Nevertheless, global abatement may increase or decrease depending on the asymmetry among countries. Even if transfers are allowed, countries may end up in a SB outcome if the degree of asymmetry among countries is not sufficiently high. Besides asymmetries between countries, domestic policies such as trade, investment and adaptation could also affect the bargaining positions of countries in climate negotiations, and hence the cooperative outcome and the allocation of abatement duties.

In this paper, we presented the basic features of the interaction among countries in mitigating climate change. We focused on the effects of asymmetries among countries on the outcomes of negotiations and the allocation of abatement burdens. We restricted our analysis to two countries, implying that the condition of individual rationality is sufficient for cooperation. Generalising the model to n countries adds a more challenging participation constraint due to free-rider incentives. This constraint is known in the literature as the 'internal stability condition'. That is, each country chooses to join an agreement if it receives at least as high as its free-rider payoff which is obtained by remaining outside the agreement while other countries cooperate. The internal stability condition is stronger than individual rationality because of the positive externalities associated with abatement efforts. As a result, as mentioned in Section 1, the IEA literature concludes that a small number of countries will cooperate. We can solve the constrained maximisation using the internal stability condition as a participation constraint and solve the NBS using the free-rider payoff as a threat point. Although the individual rationality condition is always satisfied under these two SB solutions, the internal stability condition may be violated. Therefore, we may reach the same conclusion as in the literature that only small agreements are stable under the SB solutions.

Acknowledgements

We are grateful to Prof. Alejandro Caparrós for his helpful comments, and the participants of the Second AERNA Workshop on Game Theory and the Environment, Madrid, Spain, September 2017.

References

Altamirano-Cabrera, J.-C., Finus, M., and Dellink, R. (2008). Do abatement quotas lead to a more successful climate coalition? *The Manchester School*, 51(1):93–109.

Barrett, S. (1994). Self-enforcing international environmental agreements. *Oxford Economic Papers*, 46:878–894.

Barrett, S. (1997). Heterogeneous international environmental agreements. In Carraro, C., ed., *International Environmental Agreements: Strategic Policy Issues*. Edward Elgar, Cheltenham.

Barrett, S. (2001). International cooperation for sale. *European Economic Review*, 45(10):1835–1850.

Bayramoglu, B. and Jacques, J.-F. (2015). International environmental agreements: the case of costly monetary transfers. *Environmental and Resource Economics*, 62(4):745–767.

Binmore, K. (1998). *Game Theory and the Social Contract-Vol. 2: Just Playing*. MIT press, Cambridge, UK.

Boom, J. T. (2006). *International emissions trading: design and political acceptability*. PhD thesis, University of Groningen.

Botteon, M. and Carraro, C. (1997). Burden-sharing and coalition stability in environmental negotiations with asymmetric countries. In Carraro, C., ed., *International Environmental Negotiations: Strategic Policy Issues*, pp. 26–55. Cheltenham, UK: Elgar.

Buchholz, W. and Konrad, K. A. (1994). Global environmental problems and the strategic choice of technology. *Journal of Economics*, 60(3):299–321.

Carraro, C., Eyckmans, J., and Finus, M. (2006). Optimal transfers and participation decisions in international environmental agreements. *The Review of International Organizations*, 1(4):379–396.

Carraro, C. and Siniscalco, D. (1993). Strategies for the international protection of the environment. *Journal of Public Economics*, 52(3):309–328.

Chichilnisky, G. and Heal, G. (1994). Who should abate carbon emissions? An International Viewpoint. *Economics Letters*, 44(4):443–449.

Copeland, B. R. and Taylor, M. S. (2005). *Trade and the Environment: Theory and Evidence*. Princeton university press.

d'Aspremont, C., Jacquemin, A., Gabszewicz, J. J., and Weymark, J. A. (1983). On the stability of collusive price leadership. *Canadian Journal of Economics*, 16(1):17–25.

Dellink, R., Finus, M., van Ierland, E., and Altamirano, J. (2004). Empirical background paper of the STACO model. Wageningen University.

Douven, R. C. and Engwerda, J. C. (1995). Properties of n-person axiomatic bargaining solutions if the pareto frontier is twice differentiable and strictly concave. Center of Economic Research, Tilburg University.

Endres, A. and Finus, M. (2002). Quotas may beat taxes in a global emission game. *International Tax and Public Finance*, 9(6):687–707.

Engwerda, J. (2005). *LQ Dynamic Optimization and Differential Games*. John Wiley & Sons.

Escapa, M. and Gutiérrez, M. J. (1997). Distribution of potential gains from international environmental agreements: the case of the greenhouse effect. *Journal of Environmental Economics and Management*, 33(1):1–16.

Eyckmans, J. (2009). International environmental agreements and the case of global warming. In Yew-Kwang Ng, I. W., ed., *Welfare Economics and Sustainable Development –Volume II*. Encyclopedia of Life Support Systems (EOLSS), Oxford, UK.

Eyckmans, J. and Cornillie, J. (2002). Efficiency and equity in the EU burden sharing agreement. Working Paper 2000-02, Center of Economic Studies., KU Leuven.

Eyckmans, J. and Finus, M. (2006). New roads to international environmental agreements: the case of global warming. *Environmental Economics and Policy Studies*, 7(4):391–414.

Eyckmans, J., Proost, S., and Schokkaert, E. (1993). Efficiency and distribution in greenhouse negotiations. *Kyklos*, 46(3):363–397.

Finus, M. (2001). *Game Theory and International Environmental Cooperation*. Cheltenham, UK and Northampton, MA: Edward Elgar.

Finus, M. (2008). Game theoretic research on the design of international environmental agreements: insights, critical remarks, and future challenges. *International Review of Environmental and Resource Economics*, 2(1):29–67.

Finus, M. and McGinty, M. (2019). The anti-paradox of cooperation: diversity may pay! *Journal of Economic Behavior and Organization*, 157:541–559.

Finus, M. and Rundshagen, B. (1998). Toward a positive theory of coalition formation and endogenous instrumental choice in global pollution control. *Public Choice*, 96(1-2):145–186.

Friedman, J. W. (1986). *Game Theory with Applications to Economics*. Oxford University Press, USA.

Harstad, B. (2016). The dynamics of climate agreements. *Journal of the European Economic Association*, 14(3):719–752.

Helm, D., Hepburn, C., and Ruta, G. (2012). Trade, climate change, and the political game theory of border carbon adjustments. *Oxford Review of Economic*

Policy, 28(2):368–394.

Hoel, M. (1991). Global environmental problems: the effects of unilateral actions taken by one country. *Journal of Environmental Economics and Management*, 20(1):55–70.

Hoel, M. (1992). International environment conventions: the case of uniform reductions of emissions. *Environmental and Resource Economics*, 2(2):141–159.

Höhne, N., Kuramochi, T., Warnecke, C., Röser, F., Fekete, H., Hagemann, M., Day, T., Tewari, R., Kurdziel, M., Sterl, S., et al. (2017). The Paris agreement: resolving the inconsistency between global goals and national contributions. *Climate Policy*, 17(1):16–32.

McGinty, M. (2007). International environmental agreements among asymmetric nations. *Oxford Economic Papers*, 59(1):45–62.

Nash Jr, J. F. (1950). The bargaining problem. *Econometrica: Journal of the Econometric Society*, pp. 155–162.

Pavlova, Y. and De Zeeuw, A. (2013). Asymmetries in international environmental agreements. *Environment and Development Economics*, 18(1):51–68.

Zehaie, F. (2009). The timing and strategic role of self-protection. *Environmental and Resource Economics*, 44(3):337.

Appendix

A Appendix of Section 2

A.1

If we have two countries, equation (3) can be written as $\alpha_1 W_1(q)+\alpha_2 W_2(q)$, where $q=(q_1,q_2)$. $q^* = \underset{q\in\mathbb{Q}}{\arg\max}\ \alpha_1 W_1(q)+\alpha_2 W_2(q)$. Maximisation of this function gives the first-order condition as $\alpha_1 W_1'(q)+\alpha_2 W_2'(q)=0$, at any Pareto point. This gives q^* as a continuously differentiable function of α, where $\alpha=(\alpha_1,a_2)$. The Pareto frontier can be parametrised by α_1 since $\alpha_2 = 1-\alpha_1$. Now, the outcome on the Pareto curve is determined by α_1, so we have $W_1^*(\alpha_1)$ and $W_2^*(\alpha_1)$. Substituting in the first-order condition above, we have $\alpha_1 \frac{dW_1(q^*(\alpha_1))}{d\alpha_1} + \alpha_2 \frac{dW_2(q^*(\alpha_1))}{d\alpha_1} = 0$, which gives $\frac{dW_2(q^*(\alpha_1))/d\alpha_1}{dW_1(q^*(\alpha_1))/d\alpha_1} = -\frac{\alpha_1}{\alpha_2}$. Therefore, the slope of any line tangent to the Pareto frontier is $\frac{dW_2(q^*(\alpha_1))}{dW_1(q^*(\alpha_1))} = -\frac{\alpha_1}{\alpha_2}$.[15]

[15] For the proof of n-countries, see Douven and Engwerda (1995), Appendix A, pp.15-16.

A.2 Constrained Joint Welfare Maximisation

From the Lagrangian in (11), the first-order condition is given by:

$$\begin{aligned}\frac{\partial L}{\partial q_i} &= 0 \rightarrow B_1'(Q)(1+\lambda_1) + B_2'(Q)(1+\lambda_2) + ... + B_n'(Q)(1+\lambda_n) \\ &= C_i'(q_i)(1+\lambda_i)\,. \end{aligned} \tag{A.1}$$

The Kuhn-Tucker (KT) conditions are: $\frac{\partial L}{\partial \lambda_i} = B_i(Q) - C_i(q_i) - W_i^{*NE} \geq 0$, $\lambda_i \geq 0$ and $\lambda_i \frac{\partial L}{\partial \lambda_i} = 0$.

Equation (12) can be re-written to compare the constrained with the unconstrained joint welfare maximisation as follows:

$$\sum_{j \in N} B_j'(Q) + \sum_{j \in N} \lambda_j B_j'(Q) = (1+\lambda_i) C_i'(q_i)\,. \tag{A.2}$$

Following Hoel (1992), the above equation can be simplified, where $\sum_{j \in N} B_j'(Q)$ can be written as $\sum_{j \in N} B_j'(Q) = (1+\lambda_i)\sum_{j \in N} B_j'(Q) - \lambda_i \sum_{j \in N} B_j'(Q)$, which is equal to $\sum\limits_{j \in N} B_j'(Q) + \lambda_i \sum\limits_{j \in N} B_j'(Q) - \lambda_i \sum\limits_{j \in N} B_j'(Q)$, and, hence we obtain:

$$(1+\lambda_i)\sum_{j \in N} B_j'(Q) - \lambda_i \sum_{j \in N} B_j'(Q) + \sum_{j \in N} \lambda_j B_j'(Q) = (1+\lambda_i) C_i'(q_i)\,. \tag{A.3}$$

Divide (A.3) by $1+\lambda_i$ delivers (12) in the text.

A.3 Nash Bargaining Solution

Differentiation of (17) with respect to q_i gives:

$$\begin{aligned}&\frac{B_1'(Q) - C_1'(q_1)}{B_1(Q) - C_1(q_1) - W_1^{*NE}} + \frac{B_2'(Q)}{B_2(Q) - C_2(q_2) - W_2^{*NE}} \\ &+ .. + \frac{B_n'(Q)}{B_n(Q) - C_n(q_n) - W_n^{*NE}} = 0\,, \end{aligned} \tag{A.4}$$

$$\begin{aligned}&\frac{B_1'(Q)}{B_1(Q) - C_1(q_1) - W_1^{*NE}} + \frac{B_2'(Q)}{B_2(Q) - C_2(q_2) - W_2^{*NE}} \\ &+ .. + \frac{B_n'(Q) - C_n'(q_n)}{B_n(Q) - C_n(q_n) - W_n^{*NE}} = 0\,. \end{aligned} \tag{A.5}$$

Let $B_i(Q) - C_i(q_i) - W_i^{*NE} = G_i$, where G_i is the gain from cooperation.

Therefore, the above equations can be written as:

$$\frac{B_1^{'}(Q)}{G_1} + \frac{B_2^{'}(Q)}{G_2} + ... + \frac{B_n^{'}(Q)}{G_n} = \frac{C_1^{'}(q_1)}{G_1}\,, \tag{A.6}$$

$$\frac{B_1^{'}(Q)}{G_1} + \frac{B_2^{'}(Q)}{G_2} + ... + \frac{B_n^{'}(Q)}{G_n} = \frac{C_n^{'}(q_n)}{G_n}\,, \tag{A.7}$$

which leads to:

$$\sum_{j\in N} \frac{1}{G_j} B_j^{'}(Q) = \frac{1}{G_i} C_i^{'}(q_i)\,. \tag{A.8}$$

Multiply the above equation by G_i, we obtain (18) in the text:

$$G_i \sum_{j\in N} \frac{1}{G_j} B_j^{'}(Q) = C_i^{'}(q_i)\,. \tag{A.9}$$

To compare the NBS with the non-cooperative outcome, we can re-write the above equation as:

$$B_i^{'}(Q) + G_i \sum_{j\in N, j\neq i} \frac{B_j^{'}(Q)}{G_j} = C_i^{'}(q_i)\ \ . \tag{A.10}$$

B Appendix of Section 3

B.1 Equilibrium Abatement and Welfare Levels

Using (22) and (23):

B.1.1 The Non-cooperative Outcome

If countries maximise their own welfare function, the non-cooperative (NE) abatement levels are given by:

$$q_1^{*NE} = \frac{b}{c} \ \ \& \ \ q_2^{*NE} = \frac{\beta b}{\gamma c}\,. \tag{A.11}$$

By inserting the above equilibrium abatement levels into (22) and (23), we obtain $W_1^{*NE} = \frac{1}{2}\frac{b^2(2\beta+\gamma)}{\gamma c}$ and $W_2^{*NE} = \frac{1}{2}\frac{\beta b^2(\beta+2\gamma)}{\gamma c}$, which are the welfare levels of countries if negotiations fail, i.e. at the disagreement point.

B.1.2 The FB Solution

Under the FB outcome, solving (7), leads to:

$$q_1^{*Unc} = \frac{b}{c} + \frac{\beta b}{c} = \frac{b\,(\beta+1)}{c} \;\&\; q_2^{*Unc} = \frac{\beta b}{\gamma c} + \frac{b}{\gamma c} = \frac{b\,(\beta+1)}{\gamma c}\,. \tag{A.12}$$

Consequently, we get the equilibrium welfare level of each country, $W_1^{*Unc} = \frac{1}{2}\frac{b^2(\beta+1)(2+(1-\beta)\gamma)}{\gamma c}$ and $W_2^{*Unc} = \frac{1}{2}\frac{b^2(\beta+1)(2\beta\gamma+\beta-1)}{\gamma c}$.

B.1.3 Constrained Joint Welfare Maximisation

Under the constrained outcome, countries allocate their abatement efforts by solving (10), from which the equilibrium abatement levels q_i^{*Con} follows:

$$q_1^{*Con} = \frac{b\,(\beta+1)}{c} + \frac{(\lambda_2-\lambda_1)\,\beta b}{(1+\lambda_1)\,c} \;\&\; q_2^{*Con} = \frac{b\,(\beta+1)}{\gamma c} + \frac{(\lambda_1-\lambda_2 b)}{(1+\lambda_2)\,\gamma c}\,. \tag{A.13}$$

Solving the first-order condition in (A.1) and the KT conditions, which are given by $\partial L/\partial\lambda_1 \geq 0$, $\partial L/\partial\lambda_2 \geq 0$, $\lambda_1\,(\partial L/\partial\lambda_1\,) = 0$ and $\lambda_2\,(\partial L/\partial\lambda_2\,) = 0$, there are basically four configurations possible at the optimum:

i) If constraints of both countries are not binding, i.e. $\lambda_1 = \lambda_2 = 0$, which corresponds to the unconstrained maximisation, we obtain the same equilibrium abatement levels as in (A.12). Substituting these abatement levels in $\partial L/\partial\lambda_1 > 0$, we obtain $-\frac{1}{2}\frac{b^2\left(\beta^2\gamma-2\right)}{\gamma c} > 0 \;\Leftrightarrow\;$ **a**) $\beta^2 < \frac{2}{\gamma}$, while if we substitute these abatement levels in $\partial L/\partial\lambda_2 > 0$ we get $\frac{1}{2}\frac{b^2\left(2\beta^2\gamma-1\right)}{\gamma c} > 0 \Leftrightarrow$ **b**) $\beta^2 > \frac{1}{2\gamma}$ which are the two conditions stated in Proposition 1.

ii) If the constraint of country 1 is binding, i.e. $\lambda_1 > 0$ *and* $\lambda_2 = 0$, we solve the first-order condition $\partial L/\partial q_2 = 0$ for λ_1 and obtain $\lambda_1 = \frac{\gamma c q_2 - b(1-\beta)}{b}$ Then, we substitute the value of λ_1 in the first-order condition $\partial L/\partial q_1 = 0$. With respect to the KT conditions, since $\lambda_1 > 0$, we have $\partial L/\partial\lambda_1 = 0$. Solving these two equations: $\partial L/\partial q_1 = 0$ and $\partial L/\partial\lambda_1 = 0$, we obtain $q_1^{*Con} = \frac{b\left(\left(2\beta\gamma^2\right)^{1/3}+\gamma\right)}{\gamma c}$ and $q_2^{*Con} = \frac{1}{2}\frac{b\left(\left(2\beta\gamma^2\right)^{2/3}+2\beta\gamma\right)}{\gamma^2 c}$.

iii) Similarly, if the constraint of country 2 is binding $\lambda_1 = 0$ *and* $\lambda_2 > 0$, we solve the first-order condition $\partial L/\partial q_1 = 0$ for λ_2 and obtain $\lambda_2 = \frac{cq_1 - b(1-\beta)}{b\beta}$. Solving the two equations: $\partial L/\partial q_2 = 0$ and $\partial L/\partial\lambda_2 = 0$ delivers $q_1^{*Con} = \frac{1}{2}\frac{b\left(\left(2\beta^2\gamma\right)^{2/3}+2\beta\gamma\right)}{\beta\gamma c}$ and $q_2^{*Con} = \frac{b\left(\left(2\beta^2\gamma\right)^{1/3}+\beta\right)}{\gamma c}$. The associated equilibrium welfare levels are $W_1^{*Con} =$

$\frac{1}{4}\frac{b^2\left(3\ \left(2\beta^2\gamma\right)^{1/3}+4\beta+2\gamma\right)}{\gamma c}$ and $W_2^{*Con} = W_2^{*NE}$.

iv) Finally, if we assume the constraints of both countries are binding, i.e. $\lambda_1 > 0$ *and* $\lambda_2 > 0$. By solving the two KT equations: $\partial L/\partial\lambda_1 = 0$ and $\partial L/\partial\lambda_2 = 0$, we obtain $q_1 = \frac{b}{c}$ and $q_2 = \frac{\beta b}{\gamma c}$, which are the NE abatement levels. Substituting these equilibrium abatement levels into the two first-order conditions $\partial L/\partial q_1 = 0$ and $\partial L/\partial q_2 = 0$ and solving for λ_1 and λ_2, we get $\lambda_1 = -1$ and $\lambda_2 = -1$ which contradicts the KT conditions $\lambda_i \geq 0\ \forall i = 1, 2$. Therefore, it is impossible for the two countries' constraints to be binding.

B.1.4 The Nash Bargaining Solution

If both countries allocate their abatement burdens using the NBS as shown in (16), we have the following:

$$q_1^{*NBS} = \frac{b}{c} + \frac{\beta b\frac{G_1}{G_2}}{c} = \frac{b\left(\beta\frac{G_1}{G_2} + 1\right)}{c}\ \&\ q_2^{*NBS} = \frac{\beta b}{\gamma c} + \frac{b\frac{G_2}{G_1}}{\gamma c} = \frac{b\left(\beta + \frac{G_2}{G_1}\right)}{\gamma c}. \quad \text{(A.14)}$$

which can be explicitly written as:

$$q_1^{*NBS} = \frac{b\left(\left(\beta\gamma^2\right)^{1/3} + \gamma\right)}{\gamma c}\ \&\ q_2^{*NBS} = \frac{b\left(\left(\beta\gamma^2\right)^{2/3} + \beta\gamma\right)}{\gamma^2 c}. \quad \text{(A.15)}$$

The associated equilibrium welfare levels are $W_1^{*NBS} = \frac{b\left(\left(\beta\gamma^2\right)^{1/3}+\gamma\right)}{\gamma c}$ and $W_2^{*NBS} = \frac{b\left(\left(\beta\gamma^2\right)^{2/3}+\beta\gamma\right)}{\gamma^2 c}$.

B.2 Proof of Corollary 1

i. For country 1: $C_1'\left(q_1^{*Unc}\right) = b\left(1+\beta\right)$, $C_1'\left(q_1^{*Con}\right) = b\left(1+\beta+\lambda_2\beta\right)$, and $C_1'\left(q_1^{*NBS}\right) = b\left(1+\frac{G_1}{G_2}\beta\right)$. First, since $b\left(1+\beta+\lambda_2\beta\right) > b\left(1+\beta\right)$ given $\lambda_2 > 0$, hence we have $C_1'\left(q_1^{*Con}\right) > C_1'\left(q_1^{*Unc}\right) \Longleftrightarrow q_1^{*Con} > q_1^{*Unc}$. Second, we have shown in Section 3 that the slope of the line tangent to the Pareto curve at the NBS is flatter than the constrained outcome, i.e. $\frac{(1+\lambda_1)}{(1+\lambda_2)} > \frac{G_2}{G_1}$. Hence, $\frac{G_1}{G_2} > \frac{(1+\lambda_2)}{(1+\lambda_1)} \Leftrightarrow \frac{G_1}{G_2} > (1+\lambda_2)$ given our assumption that $\lambda_1 = 0$. Therefore, by comparing the constrained maximisation with the NBS we have: $b\left(1+\frac{G_1}{G_2}\beta\right) > b\left(1+\beta\left(1+\lambda_2\right)\right)$ since $\frac{G_1}{G_2} > (1+\lambda_2)$. Thus, $C_1'\left(q_1^{*NBS}\right) > C_1'\left(q_1^{*Con}\right) \Longleftrightarrow q_1^{*NBS} > q_1^{*Con}$.

ii. For country 2: $C_2'\left(q_2^{*Unc}\right) = b\,(1+\beta)$, $C_2'\left(q_2^{*Con}\right) = b\left(1+\beta-\frac{\lambda_2}{1+\lambda_2}\right)$, and $C_2'\left(q_2^{*NBS}\right) = b\left(\beta+\frac{G_2}{G_1}\right)$. Since $b\left(1+\beta-\frac{\lambda_2}{1+\lambda_2}\right) < b\,(1+\beta)$ given $\lambda_2 > 0$, hence we have $C_2'\left(q_2^{*Con}\right) < C_2'\left(q_2^{*Unc}\right) \Longleftrightarrow q_2^{*Con} < q_2^{*Unc}$. Furthermore, $b\left(1+\beta-\frac{\lambda_2}{1+\lambda_2}\right)$ can be written as $b\left(\beta+\frac{1}{1+\lambda_2}\right)$, thus we have $b\left(\beta+\frac{G_2}{G_1}\right) < b\left(\beta+\frac{1}{1+\lambda_2}\right)$ since $\frac{G_1}{G_2} > (1+\lambda_2)$ as we have shown in the first part of this proof. Thus, $C_2'\left(q_2^{*NBS}\right) < C_2'\left(q_2^{*Con}\right) \Longleftrightarrow q_2^{*NBS} < q_2^{*Con}$.

B.3 Proof of Corollary 2

i. Lemma 1 shows that $C_1'(q_1) > C_2'(q_2)$, hence we have $cq_1^{*Con} > \gamma c q_2^{*Con}$ and $cq_1^{*NBS} > \gamma c q_2^{*NBS}$. Since both countries have the same cost functions, i.e. $\gamma = 1$, country 1 contributes more than country 2 under the constrained maximisation and the NBS.

ii. Given Assumption 1, country 2 receives no gains from cooperation under the constrained outcome. Therefore, $W_1^{*Con} - W_1^{*NE} > W_2^{*Con} - W_2^{*NE} = 0$. As country 1 has a steeper marginal cost function than country 2, i.e. $\gamma < 1$, it enjoys a higher welfare level than country 2 in the non-cooperative outcome, $W_1^{*NE} - W_2^{*NE} = \frac{1}{2}\frac{b^2(1-\gamma)}{\gamma c} > 0$. Therefore, $W_1^{*Con} > W_1^{*NE} > W_2^{*NE} = W_2^{*Con}$. To determine which country contributes more, assume on the contrary that $q_1^{*Con} > q_2^{*Con}$. In this case, country 1 incurs higher costs than country 2 while both countries enjoy the same benefits: $C_1(q_1) > C_2(q_2)$ and $B_1(Q) = B_2(Q)$. Thus, we obtain $W_1^{*Con} < W_2^{*Con}$, which is a contradiction. Thus we have $q_1^{*Con} < q_2^{*Con}$. Similarly, we have $G_1 > G_2$ and $C_1'(q_1) > C_2'(q_2)$ as shown in Lemma 1. Therefore, $G_1 = W_1^{*NBS} - W_1^{*NE} > G_2 = W_2^{*NBS} - W_2^{*NE}$. Assume that $q_1^{*NBS} > q_2^{*NBS}$. In this case, country 1 has higher costs than country 2 while both countries enjoys the same benefits. Thus, we obtain $W_1^{*NBS} < W_2^{*NBS}$. However, we have $W_1^{*NE} > W_2^{*NE}$. From the above information, $G_1 - G_2 = -\frac{1}{2}\frac{c^2\gamma\left(q_1^{2*NBS} - \gamma q_2^{2*NBS}\right) + b^2(1-\gamma)}{\gamma c} < 0$, which is a contradiction, so we have $q_1^{*NBS} < q_2^{*NBS}$.

B.4 Proof of Corollary 3

In the case of a negative $\beta - \gamma$ covariance, combining the results of the previous appendix, it is straightforward to check that the country which has a higher marginal benefits and a flatter marginal cost curve contributes more than the other country. i) If $\beta < 1$ and $\gamma > 1$, we get $q_1^{*Con} > q_2^{*Con}$ and $q_1^{*NBS} > q_2^{*NBS}$, given $cq_1^{*Con} > \gamma c q_2^{*Con}$ and $cq_1^{*NBS} > \gamma c q_2^{*NBS}$ as shown in Lemma 1. ii) If $\beta > 1$

and $\gamma < \frac{1}{2}$, and using Appendix B.1.3(iii), $q_1^{*Con} < q_2^{*Con}$ if $\beta > \frac{1}{\sqrt{2}}\gamma$, where $1 > \frac{1}{\sqrt{2}}\gamma$ for all $\gamma < \frac{1}{2}$, hence we have, $q_1^{*Con} < q_2^{*Con}$.

In the case of a positive $\beta - \gamma$ covariance, i.e. if $\beta < 1$ and $\gamma < 1$. Using Appendix B.1.3(iii) again, we have $q_1^{*Con} \geq (<)\, q_2^{*Con}$ if $\beta \leq (>)\, \frac{1}{\sqrt{2}}\gamma$. However, Assumption 1(b) is not sufficient to guarantee this condition, so it needs to hold. Similarly, for the NBS, using (A.15), $q_1^{*NBS} \geq (<)\, q_2^{*NBS}$ if $\beta \leq (<)\, \gamma$. Given Assumption 1(b), we can only obtain $\beta \leq \gamma$ if $\gamma < \overline{\gamma} \simeq 0.79$.

B.5 Proof of Corollary 4

We use the equilibrium abatement levels in Appendix B.1.

- For $\gamma = 1$, (i.e. $\beta < 0.707$, given Assumption 1(b)), we have $Q^{*Unc} > Q^{*Con}$ if $\beta < 0.707$, which always holds. Similarly, we get $Q^{*Unc} > Q^{*NBS}$ if $\beta < 1$. However, solving for $Q^{*Con} - Q^{*NBS}$, we find that $Q^{*Con} \geq (<)\, Q^{*NBS}$ if $\beta > \frac{1}{2}$.

- For $\beta = 1$, (i.e. $\gamma < 0.5$, given Assumption 1(b)). Similar to the previous case, we obtain $Q^{*Unc} > Q^{*Con}$ if $\gamma < 0.5$, and $Q^{*Unc} > Q^{*NBS}$ if $\gamma < 1$. In addition, we get $Q^{*Con} > Q^{*NBS}$ if $\gamma < 2$. Therefore, we have the following ranking of global abatement levels: $Q^{*Unc} > Q^{*Con} > Q^{*NBS}$.

- For a negative $\beta - \gamma$ covariance: a)$\beta < 1$ and $\gamma > 1$, we find that $Q^{*Unc} > Q^{*Con}$ if either $\beta > \sqrt{\frac{1}{2\gamma}}$ or $\beta < \frac{2}{\gamma^2}$. The first solution violates Assumption 1(b) and hence we ignore this solution. Therefore, $Q^{*Unc} > Q^{*Con}$ if $\beta < \frac{2}{\gamma^2}$. For all $1 < \gamma < 2$, assumption 1 is sufficient to guarantee the satisfaction of this condition, thus we have $Q^{*Unc} > Q^{*Con}$. However, if $\gamma > 2$, we need $\beta < \frac{2}{\gamma^2}$ to hold. If $\beta > \frac{2}{\gamma^2}$, we have $Q^{*Unc} < Q^{*Con}$ $\forall \gamma > 2$. Similarly, we obtain $Q^{*Unc} > Q^{*NBS}$ if $\beta < \frac{1}{\gamma^2}$, which holds if $1 < \gamma < 1.26$ such that $\beta < \sqrt{\frac{1}{2\gamma}} < \frac{1}{\gamma^2}$. However, if $\gamma > 1.26$, we need this condition to hold. Therefore, if $\beta > \frac{1}{\gamma^2}$, we have $Q^{*Unc} < Q^{*NBS}$. Finally, we have $Q^{*Con} > Q^{*NBS}$ if and only if $\beta > \frac{1}{2}\gamma$, which violates Assumption 1(b) if $\gamma > 1.26$. Therefore, for all $\gamma > 1.26$ we have $\beta < \sqrt{\frac{1}{2\gamma}} < \frac{1}{2}\gamma$, hence $Q^{*Con} < Q^{*NBS}$.

  b) $\beta > 1$ and $\gamma < 1$, we have , as shown above, $Q^{*Unc} > Q^{*Con}$ if $\beta < \frac{2}{\gamma^2}$, which holds for all values of γ given $\beta < \sqrt{\frac{1}{2\gamma}}$. In addition, we always have $Q^{*Unc} > Q^{*NBS}$ since $\beta < \sqrt{\frac{1}{2\gamma}} < \frac{1}{\gamma^2}$ for all values of $\gamma < 1$. Furthermore, given $\gamma < 1$ and $\beta > 1$, we always have in this case $\beta > \frac{1}{2}\gamma$, hence $Q^{*Con} > Q^{*NBS}$.

- For a positive $\beta - \gamma$ covariance, i.e. $\beta < 1$ *and* $\gamma < 1$. Since we have $\gamma < 1$, therefore, $\beta < \sqrt{\frac{1}{2\gamma}} < \frac{2}{\gamma^2}$, which leads to $Q^{*Unc} > Q^{*Con}$. Similarly, we have $\beta < \sqrt{\frac{1}{2\gamma}} < \frac{1}{\gamma^2}$, which leads to $Q^{*Unc} > Q^{*NBS}$. However, we have $\frac{1}{2}\gamma < \sqrt{\frac{1}{2\gamma}}$, therefore, $Q^{*Con} > Q^{*NBS}$ if and only if $\beta > \frac{1}{2}\gamma$, while if $\beta > \frac{1}{2}\gamma$, we have $Q^{*Con} < Q^{*NBS}$.

B.6 The Closing the Gap Index

Using Appendix B.1 and (24), we obtain $CGI^{Con} = \frac{3}{2}\frac{2^{1/3}\beta^{2/3}\gamma^{1/3}}{\beta^2\gamma+1} \cdot 100$ and $CGI^{NBS} = \frac{\beta^{1/3}(\beta^{1/3}\gamma^{2/3}+\beta\gamma}{\gamma^{1/3}(\beta^2\gamma+1)} \cdot 100$. For the case of benefit-asymmetry, we have $\frac{\partial CGI^{Con}}{\partial\beta} > 0 \ \forall\beta \in \left(0, \sqrt{\frac{1}{2}}\right)$ and $\frac{\partial CGI^{NBS}}{\partial\beta} > 0 \forall\beta \in (0,1)$ given $\gamma = 1$. Similarly, for the cost-asymmetry case, $\frac{\partial CGI^{Con}}{\partial\gamma} > 0 \ \forall\gamma \in \left(0, \frac{1}{2}\right)$ and $\frac{\partial CGI^{NBS}}{\partial\gamma} > 0\forall\gamma \in (0,1)$ given $\beta = 1$.

B.7 Proof of Corollary 5

We use the equilibrium welfare levels in Appendix B.1 and $W_i^{*Unc+\tau}$ as defined in (25).

- For $\gamma = 1$, we have $W_1^{*Unc+\tau} - W_1^{*NBS} = \frac{1}{4}\frac{b^2\left(\beta^2-2\beta^{2/3}+1\right)}{c} \geq 0$ if $\beta \leq 0.485$ or $\beta \geq 1$. Given Assumption 1(b), $\beta < 0.707$ if $\gamma = 1$. Similarly, $W_1^{*Unc+\tau} - W_1^{*Con} = -\frac{1}{4}\frac{b^2\left(3\left(2\beta^2\right)^{1/3}-\beta^2-1\right)}{c} \geq 0$ if $\beta \leq 0.14$ or $\beta \geq 2.4$, where the second condition violates Assumption 1(b). Therefore, as long as $\beta < 0.14$, we have $W_1^{*Unc+\tau} > W_1^{*Con} > W_1^{*NBS}$.

- For $\beta = 1$, we find that $W_1^{*Unc+\tau} - W_1^{*NBS} = \frac{1}{4}\frac{b^2\left(\gamma^2-2\left(\gamma^2\right)^{2/3}+\gamma\right)}{\gamma^2 c} \geq 0$ if $\gamma \leq 0.236$ or $\gamma \geq 1$. However, given Assumption 1(b), $\gamma < \frac{1}{2}$ if $\beta = 1$. Similarly, we obtain $W_1^{*Unc+\tau} - W_1^{*Con} = -\frac{1}{4}\frac{b^2\left(32^{1/3}\gamma^{1/3}-\gamma-1\right)}{\gamma c} \geq 0$ if $\gamma \leq 0.019$ or $\gamma \geq 5.8$, where the second condition does not hold. Therefore, as long as $\gamma < 0.019$, we find $W_1^{*Unc+\tau} > W_1^{*Con} > W_1^{*NBS}$.

Part III

Essay 2: Non-Cooperative Climate Policies among Asymmetric Countries: Production- versus Consumption-based Carbon Taxes

Appendix 6B: Statement of Authorship

This declaration concerns the article entitled:			
Non-Cooperative Climate Policies among Asymmetric Countries: Production- versus Consumption-based Carbon Taxes			
Publication status (tick one) **Draft manuscript** X **Submitted** ☐ **In review** ☐ **Accepted** ☐ **Published** ☐			
Publication details (reference)			
Copyright status (tick the appropriate statement) I hold the copyright for this material X Copyright is retained by the publisher, but I have been given permission to replicate the material here ☐			
Candidate's contribution to the paper (provide details, and also indicate as a percentage)	The candidate contributed to / considerably contributed to / predominantly executed the… Formulation of ideas: - Predominantly contributed to the formulation of ideas. (80 %) Design of methodology: - Predominantly contributed to the design of methodology. (90 %) ~~Experimental work:~~ Presentation of data in journal format: - Predominantly contributed to the presentation of data in journal format. (80%)		
Statement from Candidate	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature.		
Signed	Noha Nagi Elboghdadly	**Date**	4/10/2019

Last update: Feb 2019

Non-Cooperative Climate Policies among Asymmetric Countries: Production- versus Consumption-based Carbon Taxes

Noha Elboghdadly* and Michael Finus†

Abstract

In the absence of global action, non-cooperative production-based carbon taxes might be set inefficiently low due the concern of governments about carbon leakage and the loss of competitiveness of their emission-intensive industries. In a strategic trade model, we study the effect of a gradual shift from bilateral production- to unilateral or bilateral consumption-based carbon taxes, considering various forms of border carbon adjustments (BCAs). We analyse the optimal response of two asymmetric countries in a non-cooperative policy game. We show that if the environmentally more concerned government shifts unilaterally to a consumption-based policy, BCAs on imports create a new incentive for the optimal tax structure of both governments to tax emissions. Although profit-shifting and carbon leakage distortions are only eliminated by combining carbon tariffs with a full export rebate, the optimal tax may still be below individual marginal damages due to the strategic interaction between governments. In contrast, a bilateral consumption-based tax could be set equal to or even above individual marginal damages. We find that, in equilibrium, all forms of BCAs could allow both governments to set high carbon taxes than under a bilateral production-based tax regime. However, adding export rebates weakens the positive strategic impact of carbon tariffs on the climate policy level of the environmentally less concerned government, which suggests that BCAs with export rebates should be chosen carefully.

Keywords: Carbon Taxes, Border Carbon Adjustments, Import Tariffs, Export Rebates, Profit-shifting Effects, Tariff-shifting Effects, Leakage-shifting Effects, Price-spillover Effects.

JEL-Classification: C72, F12, F18, H23,Q58

*Department of Economics, University of Bath, 3 East, Bath, BA2 7AY, UK. Email: n.m.w.elboghdadly@bath.ac.uk

†Department of Economics, University of Graz, Universitätsstraße 15, 8010 Graz, Austria and University of Bath, 3 East, Bath, BA2 7AY, UK. Email: michael.finus@uni-graz.at

1 Introduction

Actions to mitigate climate change have mainly focused on emissions released within national boundaries of countries, referred to as the production-based approach. However, in the absence of a cooperative climate policy, this raises two main concerns. First, an environmental concern about the effectiveness of unilateral or sub-global actions: emission reductions by some environmentally friendly countries are partly or completely offset by higher emissions in other countries, a phenomenon known as 'carbon leakage'. Second, a competitiveness concern about the loss of market shares of firms located in countries with stricter climate policies. These concerns greatly affect the incentives of governments when designing their non-cooperative climate policies.

In order to address these concerns and to support a more ambitious climate policy, some economists have argued in favour of taxing emissions based on consumption rather than production (Helm et al., 2012; Peters and Hertwich, 2008; Steininger et al., 2014; Stiglitz, 2006).[1] Consumption-based climate policies can be implemented through trade instruments such as border carbon adjustments (BCAs). That is, countries complement their production-based carbon tax with carbon tariffs (BCAs on imports) and/or export rebates (BCAs on exports) in order to make up for and protect against lower carbon taxes abroad. If a full adjustment on imports is combined with a full export rebate, sometimes referred to as full BCAs, this constitutes de facto a unilateral consumption-based tax.

The strategic choice of environmental policies has been widely studied in the environmental-trade policy literature. Many of those models build on the strategic imperfect-competition trade model due to Brander and Spencer (1985), extending their model by including consumers and environmental damages. For instance, Barrett (1994), Conrad (1993) and Kennedy (1994) conclude that if environmental policy is the only instrument available to governments (i.e., a production-based environmental policy), environmental policies may be set inefficiently lax, i.e., below not only global but also individual marginal damages.

In this paper, we study whether a gradual shift from production-based to consumption-based policies, using various forms of BCAs including carbon tariffs and export rebates, could restore the effectiveness of non-cooperative climate policies. For this, we derive the general optimal tax structure for two asymmetric countries which perceive global damages from greenhouse gases differently,

[1] See also Jakob et al. (2014) for a survey of the literature that compares production with consumption-based approaches.

considering different policy regimes.

We solve a two-stage game in which governments choose their carbon taxes in the first stage, and then firms choose their output levels in the second stage. Our analysis starts with a bilateral production-based tax (PB) regime under which both governments impose a carbon tax on the home firm only. BCAs are then introduced as unilateral measures supplementing the production-based tax of the environmentally more concerned country to which the other country will respond by choosing its production-based tax. The three BCA-regimes which we consider are: (i) BCAs on imports, which fully adjust the difference between the two national tax levels, (ii) BCAs on imports and exports where the export rebate rate is chosen optimally and (iii) BCAs on imports with a full export rebate, which is de facto a unilateral consumption-based tax. Finally, we consider a bilateral consumption-based tax (CB) regime under which both governments impose a carbon tax on consumption only.

On the one hand, we show that for the country which imposes BCAs, moving gradually to a unilateral consumption-based tax mitigates both, the carbon leakage and profit-shifting effect. Only full BCAs could allow this government to fully internalise its own damages from global emissions. Nevertheless, due to the strategic interaction with the other government, the optimal tax could still be set below individual marginal damages. On the other hand, we find that the optimal tax structure of the country on which BCAs are imposed changes with carbon tariffs which incentivise it to set higher carbon taxes. This is to avoid that tax revenues go abroad. Moreover, unlike the PB-regime, carbon taxes under the CB-regime may be set equal to or above individual marginal damages, depending on the demand curvature. Even though all BCA-measures may support more ambitious carbon taxes in the two countries, adding export rebates reinforces (weakens) the effect of carbon tariffs on the climate policy level of the environmentally more (less) concerned country.

Our paper builds on the literature on strategic environmental-trade policies. This literature demonstrates that two strategic effects reduce the incentive of countries to set environmental policy at ambitious levels. First, if firms engage in Cournot competition, governments have an incentive to set lax policies to provide their firms with a strategic advantage over their rivals, which is known as the profit-shifting incentive (Barrett, 1994; Conrad, 1993; Kennedy, 1994). Second, if pollution is transboundary, the effectiveness of environmental policies is undermined by carbon leakage (Conrad, 1993; Duval and Hamilton, 2002; Kennedy, 1994). Hence, we extend this literature to show how these strategic effects could change under different policy regimes, and investigate whether new strategic effects would emerge.

Our paper benefits from the literature which compares bilateral production-based and bilateral consumption-based tax regimes for international traded products under imperfect competition, as, for instance, Haufler et al. (2005) and Haufler and Pflüger (2007). However, these studies assume symmetric countries and do not consider environmental externalities.

Moreover, our paper is related to three strands of literature, studying the effects of BCA measures on carbon leakage and/or equilibrium climate policies. The first strand focuses on quantifying the impact of BCAs on reducing leakage effects and the loss of competitiveness of emission-intensive trade-exposed (EITE) industries. Most of this strand of the literature conducts numerical analyses, based on partial or general computable equilibrium models (Babiker and Rutherford, 2005; Böhringer et al., 2012; Branger and Quirion, 2014; Fischer and Fox, 2012). They conclude that BCA measures effectively mitigate carbon leakage and the output loss of firms. They provide mixed evidence about the importance of adding export rebates to carbon tariffs (Böhringer et al., 2012; 2014; Branger and Quirion, 2014). Those models consider only unilateral climate policies. In addition, due to the complicated nature of these CGE-models, these studies assume exogenous climate policy levels and hence cannot capture the strategic role of BCAs in an endogenous policy setting as we do.

The second strand of literature also considers unilateral climate polices in theoretical models. The economic rationale for using trade measures to internalise foreign emissions goes back to Markusen (1975), who shows that a combination of a Pigouvian tax and import tariffs is optimal for the active and environmentally concerned country.[2] Some recent papers have also compared a unilateral production-based with a consumption-based carbon tax, where the latter is implemented by full BCAs. However, unlike numerical models, some papers, as for instance Jakob et al. (2013) and Eichner and Pethig (2015), conclude that full BCAs do not necessarily lead to less carbon leakage. Jakob et al. (2013) show that a consumption-based carbon tax may raise foreign emissions if the foreign country shifts its production to a more carbon intensive non-exporting sector. In a different context, assuming an intertemporal model, Eichner and Pethig (2015) find that total emissions could increase if the income effects associated with a consumption-based tax are strong. In contrast, Yomogida and Tarui (2013) show that a carbon tax with full BCAs necessarily reduces total emissions. Whereas global emissions may increase without BCAs due to strong leakage effects if the

[2]See for instance Hoel (1996) and Copeland (1996).

home firm has a sufficiently cleaner technology than the foreign firm. In such cases, full BCAs lead to a higher unilateral optimal carbon tax level.[3] Although some of these studies show the effect of BCA-measures on the policy levels of the home country, also these models do not allow for endogenous bilateral policy choices.

Our paper is closely related to the third strand which also builds on theoretical models but analyses the effect of BCA-measures in the context of bilateral endogenous policy choices. For instance, Sanctuary (2018) assumes perfect competition and shows that the exporting country may react by setting a higher carbon tax if faced with BCAs on imports. Likewise, though in an imperfect competition model, Eyland and Zaccour (2012; 2014) show, based on numerical simulations, that adding partial BCAs on imports lead both countries to set higher taxes on emissions.[4] None of the above studies consider export rebates which have a less obvious strategic effect than carbon tariffs. An exception is Hecht and Peters (2018). They show, in an imperfect competition trade model, that BCA-measures allow the home country to impose a higher carbon tax but the foreign country responds by a lower tax level. They consider BCAs on imports and symmetric BCAs on imports and exports.[5] Therefore, unlike Eyland and Zaccour, their study shows that carbon tariffs can have a negative strategic impact on the policy level of the foreign country. This result also contradicts the argument put forward by Helm et al. (2012), who conclude, albeit in a highly stylized political game without micro-foundation, that the country faced with carbon tariffs responds by setting a carbon tax on its exports. The upshot of these studies is that the role of BCAs on the climate policies of the foreign country remains inconclusive. Our paper provide a more general framework to study the incentives of governments when choosing their optimal climate policy under different BCA-regimes. In addition, we consider two forms of export rebates: an optimal and a full rebate. Hence, unlike Hecht and Peters (2018), we allow BCAs on imports and BCAs on exports to be applied at different rates.

The remainder of the paper is organised as follows. In Section 2, we present the model. In Section 3 we derive the general structure of the optimal carbon tax in the social optimum and in Section 4 we do so for the non-cooperative policy regimes. In Section 5, we use specific functions in order to rank equilibrium carbon

[3]Nicolaï et al. (2010) show that both tariffs and subsidies raise the unilateral emission tax of the home country, however, both measures are not necessarily equal to the carbon tax, hence are not exactly reflecting the idea of BCAs.

[4]In both studies, the authors do not consider the full adjustment on imports.

[5]They assume that both measures are imposed at the same rate.

taxes across different regimes. Section 6 concludes and discusses possible future research.

2 Model

We consider two countries, $i = 1, 2$, which interact in a strategic trade model, which is an extended version of Brander and Spencer (1985). There are two firms, $k = 1, 2$, where firm 1 is located in country 1, and firm 2 is located in country 2. Firms are producing a homogeneous emission-intensive good x, which generates greenhouse gas emissions. Each firm supplies the home and the foreign market; both firms compete in outputs. The inverse demand in each country is given by:

$$p_i\left(X_i\right) = u_i^{'}\left(X_i\right), \tag{1}$$

where p_i is the market price and $u_i\left(X_i\right)$ represents the utility from consuming the emission-intensive good with $u_i^{'} > 0$ and $u_i^{''} < 0$. Hence, $p_i^{'}\left(X_i\right) < 0$. Total consumption in country i is $X_i = x_{1i} + x_{2i}$, where x_{1i} and x_{2i} are the outputs supplied by firm 1 and 2 to market i, respectively, and, hence, total production is $X = X_1 + X_2$. We assume the two firms are identical and face a linear production cost function, i.e., $C_{ki}(x_{ki}) = cx_{ki}$ for $k = 1, 2$ and $i = 1, 2$.

We solve a two-stage game. In the first stage, governments simultaneously choose their climate policy levels to regulate emissions. We consider different policy regimes. In the second stage, firms simultaneously choose their output levels. The game is solved by backward induction.

2.1 Second Stage

In this stage, firms choose their profit-maximising output levels for each market. Markets are segmented. That is, firms make separate quantity decisions for the home and the foreign market. Then, profits obtained in market 1 and market 2 are given by:

$$Market\,1 : \pi_{11} = (p_1(X_1) - c - t_{11})x_{11} \;\&\; \pi_{21} = (p_1(X_1) - c - t_{21})x_{21}\,, \tag{2}$$

$$Market\,2 : \pi_{12} = (p_2(X_2) - c - t_{12})x_{12} \;\&\; \pi_{22} = (p_2(X_2) - c - t_{22})x_{22}\,, \tag{3}$$

where t_{11} (t_{21}) is the effective carbon tax which firm 1 (2) faces on its supply to market 1 and t_{12} (t_{22}) is the effective carbon tax which firm 1 (2) faces on its supply to market 2; $X_1 = x_{11} + x_{21}$ and $X_2 = x_{12} + x_{22}$. We assume a constant emission-output ratio across firms, which we normalise to 1 for simplicity, such that an emission tax is de facto an output tax.

Firms maximise their profits in (2) and (3) simultaneously, which leads to the following first-order conditions:

$$Market\,1: \frac{\partial \pi_{11}}{\partial x_{11}} = p_1 + x_{11}p_1^{'} - c - t_{11} = 0 \;\&\; \frac{\partial \pi_{21}}{\partial x_{21}} = p_1 + x_{21}p_1^{'} - c - t_{21} = 0\,, \quad (4)$$

$$Market\,2: \frac{\partial \pi_{12}}{\partial x_{12}} = p_2 + x_{12}p_2^{'} - c - t_{12} = 0 \;\&\; \frac{\partial \pi_{22}}{\partial x_{22}} = p_2 + x_{22}p_2^{'} - c - t_{22} = 0\,. \quad (5)$$

We assume that goods are strategic substitutes in market 1 and market 2. A sufficient condition to ensure this in market 1 is $\frac{\partial^2 \pi_{11}}{\partial x_{11} \partial x_{21}} = p_1^{'} + x_{11}p_1^{''} < 0$ and $\frac{\partial^2 \pi_{21}}{\partial x_{21} \partial x_{11}} = p_1^{'} + x_{21}p_1^{''} < 0$, and a similar condition can be derived for market 2. That is, reaction functions of firms are negatively sloped: increasing the production level of one firm reduces the marginal revenues of its rival which reacts by decreasing production. Assuming that the second-order conditions hold, as derived in Appendix A.1, and solving the first-order conditions in (4) ((5)) simultaneously, yield the profit-maximising output levels produced of firm 1 and 2 in both markets. The profit-maximising output levels are function of the effective taxes, i.e., $x_{1i} = f\left(t_{1i}, t_{2i}\right)$ and $x_{2i} = g\left(t_{1i}, t_{2i}\right)$. Thus, the outcome of the second stage of the game is a Nash equilibrium in output levels in each of the two markets.[6]

2.2 First Stage

Given the equilibrium output levels of firms, governments choose simultaneously the level of their carbon tax t_i in the first stage by maximising their individual welfare function under different policy regimes. We consider five non-cooperative regimes: a bilateral production-based tax (PB-regime) and a bilateral consumption-based tax (CB-regime) regime as well as three regimes under which country 1 imposes unilaterally border carbon adjustments (BCAs), supplementing its production-based tax with tariffs and export rebates, while country 2 reacts by a production-based tax (BCA-regimes).

[6] The conditions derived in Appendix A.1 ensures the existence and uniqueness of a Nash equilibrium.

The welfare function of country 1 and country 2 are given by:

$$W_1 = CS_1 + PS_1 + TR_1 - D_1(e) + BCAI_1 - BCAE_1\,, \tag{6}$$

$$W_2 = CS_2 + PS_2 + TR_2 - D_2(e)\,, \tag{7}$$

where CS_i is the consumer surplus in country i, with $CS_i = \int_0^{X_i} p_i\left(X_i\right) dX_i - p_i X_i$ which follows from (1), recalling that the total supply to market i is given by $X_i = x_{1i} + x_{2i}$. PS_i is the producer surplus, which is the total profits of the home firm k, i.e., $PS_i = \Pi_k = \pi_{k1} + \pi_{k2}$. TR_i is the tax revenue of government i, where $TR_i = t_i\left(x_{k1} + x_{k2}\right)$ under all regimes, except under the CB-regime where tax revenues are given by $TR_i = t_i\left(x_{1i} + x_{2i}\right)$.

D_i are individual damages from global greenhouse gas emissions released in the production of good x. Global damages from global emissions are $D\left(e\right)$, $e = e_1 + e_2$ where $e_1 = x_{11} + x_{12}$ and $e_2 = x_{22} + x_{21}$. That is, as we normalise the emission-output ratio to 1, global emissions are equal to total production which is equal to total consumption, i.e., $e = X$. Countries may perceive or evaluate global damages differently. Therefore, the damage function faced by country 1 and country 2 is given by:

$$D_1\left(e\right) = \gamma D\left(e\right), D_2\left(e\right) = \left(1-\gamma\right) D\left(e\right) \;\; \gamma \in \left[0.5, 1\right], \tag{8}$$

where $D^{'}\left(e\right) > 0$ and $D^{''}\left(e\right) \geq 0$. Because $\gamma \in [0.5, 1]$, country 1 is at least as concerned as country 2 about environmental damages, and, usually, more whenever γ is strictly larger than 0.5.

Finally, the last two terms in the welfare function (6), $BCAI_1$ and $BCAE_1$, stand for the tariff revenues from a unilateral BCA-policy on imports and the expenses from a tax rebate on exports, respectively. These two terms are only relevant under the three non-cooperative BCA-regimes and are zero by assumption under the PB- and the CB-regime. Given $\gamma \in [0.5, 1]$, we assume that it is country 1 which implements a unilateral BCA-policy.

In the following, we explain the difference between the five non-cooperative regimes and their impacts on effective carbon taxes which are summarised in Table 1.

First, we consider that both governments impose a production-based carbon tax on their home firm (PB-regime). Hence, the effective tax which each firm faces is equal to the tax imposed in its home country, i.e., $t_{11} = t_{12} = t_1$ and $t_{22} = t_{21} = t_2$.

Second, we consider border carbon adjustments on imports (BI-regime). Country 1 imposes not only a production-based tax but also a carbon tariff on imports from country 2.[7] Hence, firm 1 faces effective tax $t_{1i} = t_1$ as under the PB-regime, and also firm 2 faces $t_{22} = t_2$ on its supply to country 2, but faces $t_{21} = t_2 + \omega(t_1 - t_2)$ on its supply to country 1 if $t_1 > t_2$, with ω the border tax adjustment parameter on imports (Eyland and Zaccour, 2012). We assume that country 1 imposes a carbon tariff which fully adjusts the difference between the two national tax levels, i.e., $\omega = 1$. On the one hand, any value of ω above 1 would violate the equal treatment rules under the World Trade Organization (WTO).[8] On the other hand, any value of ω below 1 would not be optimal for country 1.[9] This assumption implies that both firms supplying market 1 face the same effective carbon tax. Therefore, the term $BCAI_1$ in (6) is given by $BCAI_1 = (t_1 - t_2)(x_{21})$ if $t_1 > t_2$, otherwise $BCAI_1 = 0$.

Table 1: Effective Carbon Taxes under Non-Cooperative Regimes

	Effective taxes	PB	BI	BIE	BF	CB
Market 1	t_{11}	t_1	t_1	t_1	t_1	t_1
	t_{21}	t_2	t_1	t_1	t_1	t_1
Market 2	t_{12}	t_1	t_1	$t_1(1-\hat{\varphi})$	0	t_2
	t_{22}	t_2	t_2	t_2	t_2	t_2

Third, we consider border carbon adjustments on imports and exports (BIE-regime). Under this regime, country 1 complements its production-based tax and carbon tariff with a rebate on exports to country 2. Hence, compared to the previous regimes, only the effective tax firm 1 faces on its supply to country 2 will change, which is given now by $t_{12} = t_1(1 - \hat{\varphi})$ if $t_1 > t_2$, with φ the border tax adjustment parameter on exports, which may also be called the export rebate rate.[10] $\hat{\varphi}$ indicates that φ is chosen optimally. Unlike carbon tariffs, it is not always optimal for country 1 to provide a full border rebate or adjustment to its firm on its exports. However, in order to be also consistent with the WTO equal treatment rule, we need to impose a constraint such that the effective tax firm 1 faces on its exports is at least as high as the tax of the foreign firm 2, i.e., $t_1(1 - \hat{\varphi}) \geq t_2$. This implies setting a constraint on the maximum optimal rebate

[7] We use the terms carbon tariff and BCAs on imports interchangeably.

[8] The General Agreement on Tariff and Trade (GATT) allows WTO members to apply a border tax adjustment at a rate which is not higher than the rate applied to domestically produced "like" products.

[9] See Hecht and Peters (2018) and Weitzel et al. (2012) on this point. In other words, if ω could be chosen optimally by country 1, it would choose a value above 1.

[10] Note that we cannot model BCAs on exports in the same way as on imports. That is, we cannot assume for instance $t_{12} = t_1 - \varphi(t_1 - t_2)$ as this would imply that country 2, with tax t_2, imposes a tax on output of firm 1 for market 2. However, country 2 only reacts to BCA-measures by setting a production-based tax.

rate, which will be explained in detail in Section 4. Therefore, the term $BCAE_1$ in (6) is given by $BCAE_1 = \hat{\varphi}\, t_1\, x_{12}$ if $t_1 > t_2$ and $t_1(1-\hat{\varphi}) \geq t_2$.

Fourth, we consider border carbon adjustments on imports with a full export rebate (BF-regime). That is, this regime is similar to the previous regime, except that we assume that φ is not chosen optimally but $\varphi = 1$ always. Therefore, firm 1 faces effective tax $t_{12} = 0$ on its supply to country 2. The BF-regime implies de facto a unilateral consumption-based tax imposed by country 1. Again, the WTO equal treatment rule requires that this effective tax is at least as high as t_2.[11] Because we study a bilateral endogenous policy choice, full export rebates and full adjustments on exports are not always equivalent. This is evident by considering the case where country 2 chooses a subsidy, i.e., $t_2 < 0$, in which case the fully export rebate is less than full adjustment.

Finally, we consider that both governments impose a consumption-based carbon tax (CB-regime). Hence, both firms supplying country i face t_i, i.e., $t_{11} = t_{21} = t_1$ and $t_{22} = t_{12} = t_2$. It is interesting to note that this regime is equivalent to a regime in which each country imposes a production-based tax supplemented with a tariff on imports and a full export rebate. With reference to our border adjustment parameters introduced above, this implies $\omega_i = 1$ and $\varphi_i = 1\ \forall\, i = 1, 2$. In other words, the CB-regime is equivalent to a bilateral BF-regime.

3 Optimal Climate Policy: Normative Benchmark

Before turning to the non-cooperative regimes, we briefly discuss the normative benchmark of the social optimum. We assume for simplicity of exposition that governments choose a uniform tax t_S. Hence, the Nash equilibrium output levels in the second stage are given by $x^S_{1i} = x^S_{2i} = f\,(t_S, t_S) = g\,(t_S, t_S)$.

In the first stage, governments choose the carbon tax t_S by maximising the aggregate welfare, $W = W_1 + W_2$.[12] Assuming that the second-order condition is satisfied, the socially optimal carbon tax, which is derived in Appendix A.2, can

[11] In our strategic context, this regime is compatible with the WTO articles if $t_2 \leq 0$, which is likely to be the case as we show in our example in Section 5.

[12] Both $BCAI_1$ and $BCAE_1$ in (6) are not relevant under the cooperative solution.

be written as follows:

$$\hat{t}_S = \underbrace{\frac{2p_i^{'}x_{ki}\frac{dX^S}{dt_S}}{\Lambda^S}}_{CSE\ (-)} + \underbrace{\frac{-p_i^{'}x_{ki}\frac{dX^S}{dt_S}}{\Lambda^S}}_{PSE(+)} + \underbrace{\frac{(\gamma + (1-\gamma))\ D^{'}\left(\frac{\partial e^S}{\partial t_S}\right)}{\Lambda^S}}_{EDE(+)}$$

$$\Longleftrightarrow \hat{t}_S = p_i^{'}x_{ki} + D^{'}\,, \tag{9}$$

where $D^{'} = \frac{\partial D(e)}{\partial e} > 0$, $EDE = D^{'}$, $\Lambda^S = \frac{dX^S}{dt_S} = \frac{2p_1^{'}}{J_1} + \frac{2p_2^{'}}{J_2} < 0$, and $\frac{dx_{1i}^S}{dt_S} = \frac{dx_{2i}^S}{dt_S} = \frac{p_i^{'}}{J_i} < 0$ with $J_i > 0$, which is the Jacobian matrix of the second-order conditions in the second stage. We may recall $p_i^{'} < 0$ as the demand function is downward sloping and note that $\frac{\partial e^S}{\partial t_S} = \frac{dX^S}{dt_S}$ as total emissions are equal to total output.

In this but also in the following section, the structure of the optimal climate policy level is broken down into several effects: 1) the consumer surplus effect (CSE), which is related to distortions created by underproduction associated with imperfect competition (Barnett, 1980), 2) the producer surplus effect (PSE), which is related to the profits of firms, and 3) the environmental damage effect (EDE) stemming from the internalisation of damages from global emissions. The PSE is a net effect which also considers the tax revenues of governments, i.e., $PS_i + TR_i$ which are net profits. This is because taxes paid by firms are equal to tax revenues.

In (9), the CSE is negative, hence would call for a subsidy in order to raise production in both countries. The PSE is positive and hence would call for a tax. That is, when governments choose their taxes cooperatively, imposing a positive tax will raise the collective net profits of the two firms by enforcing a monopolistic output. Finally, governments jointly internalise global damages from emissions. Hence, the EDE is positive and would call for a tax equal to global marginal damages. Although the CSE and the PSE have opposite signs, the CSE dominates. Therefore, the socially optimal carbon tax level is smaller than global marginal damages, $\hat{t}_S < D^{'}$.[13] Whether the equilibrium tax will be positive or negative cannot be deduced at this level of generality, though it increases with the value of global marginal damages.

[13] See also Kennedy (1994) and Duval and Hamilton (2002). In Conrad (1993), the cooperative tax level is larger than global marginal damages because he assumes consumption takes place in a third market and hence *CSE=0*.

4 Non-cooperative Optimal Climate Policies

In this section, we analyse the impact of a gradual shift from a bilateral production-based to a bilateral consumption-based tax on the three effects derived in the previous section. Moreover, we investigate whether new effects emerge.

4.1 Bilateral Production-based Tax (PB-regime)

We start our analysis by considering the bilateral production-based tax regime (PB-regime), which serves as reference for the later regimes. Both governments impose a carbon tax on the production of their home firm. With reference to Table 1, solving the second stage delivers: $x_{1i}^{PB} = f(t_1, t_2)$ and $x_{2i}^{PB} = g(t_1, t_2)$ for $i = 1, 2$. We denote a country's welfare function under the PB-regime by W_i^{PB}. All details are provided in Appendix A.3.

Maximising W_i^{PB} with respect to the national tax level t_i gives the optimal PB-tax structure of country 1 and country 2, respectively:[14]

$$\hat{t}_1^{PB} = \underbrace{\frac{p_1^{'}(x_{11}+x_{21})\frac{p_1^{'}}{J_1}}{\Lambda_1^{PB}}}_{CSE_1\,(-)} + \underbrace{\frac{-p_1^{'}x_{11}\frac{dx_{21}^{PB}}{dt_1} - p_2^{'}x_{12}\frac{dx_{22}^{PB}}{dt_1}}{\Lambda_1^{PB}}}_{PSE_1(-,-)} + \underbrace{\frac{\gamma\, D^{'}\left(\frac{\partial e^{PB}}{\partial t_1}\right)}{\Lambda_1^{PB}}}_{EDE_1\,(+)}, \tag{10}$$

where $\Lambda_1^{PB} = \frac{dx_{11}^{PB}}{dt_1} + \frac{dx_{12}^{PB}}{dt_1} < 0$, $\frac{\partial e^{PB}}{\partial t_1} = \frac{\partial e_1^{PB}}{\partial t_1} + \frac{\partial e_2^{PB}}{\partial t_1} < 0$, with $\underbrace{\frac{\partial e_1^{PB}}{\partial t_1}}_{HEE_1} < 0$, $\underbrace{\frac{\partial e_2^{PB}}{\partial t_1}}_{FEE_1} > 0$, $EDE_1 = \gamma\, D^{'} + \underbrace{\frac{\gamma\, D^{'}(FEE_1)}{\Lambda_1^{PB}}}_{(-)}$ and

$$\begin{aligned} \frac{dx_{11}^{PB}}{dt_1} &= \frac{2p_1^{'} + p_1^{''}x_{21}}{J_1} < 0\,, \quad \frac{dx_{21}^{PB}}{dt_1} = -\frac{p_1^{'} + p_1^{''}x_{21}}{J_1} > 0\,, \\ \frac{dx_{12}^{PB}}{dt_1} &= \frac{2p_2^{'} + p_2^{''}x_{22}}{J_2} < 0\,, \quad \frac{dx_{22}^{PB}}{dt_1} = -\frac{p_2^{'} + p_2^{''}x_{22}}{J_2} > 0\,. \end{aligned} \tag{11}$$

$$\hat{t}_2^{PB} = \underbrace{\frac{p_2^{'}(x_{22}+x_{12})\frac{p_2^{'}}{J_2}}{\Lambda_2^{PB}}}_{CSE_2\,(-)} + \underbrace{\frac{-p_1^{'}x_{21}\frac{dx_{11}^{PB}}{dt_2} - p_2^{'}x_{22}\frac{dx_{12}^{PB}}{dt_2}}{\Lambda_2^{PB}}}_{PSE_2(-,-)} + \underbrace{\frac{(1-\gamma)\, D^{'}\left(\frac{\partial e^{PB}}{\partial t_2}\right)}{\Lambda_2^{PB}}}_{EDE_2\,(+)}, \tag{12}$$

[14] The optimal taxes display the various effects that play a role when choosing taxes. They are different from equilibrium taxes which would require to solve (10) and (12) simultaneously, which is not possible at this level of generality. The same applies to all subsequent regimes.

where $\Lambda_2^{PB} = \frac{dx_{22}^{PB}}{dt_2} + \frac{dx_{21}^{PB}}{dt_2} < 0$, $\frac{\partial e^{PB}}{\partial t_2} = \frac{\partial e_1^{PB}}{\partial t_2} + \frac{\partial e_2^{PB}}{\partial t_2} < 0$, with $\underbrace{\frac{\partial e_2^{PB}}{\partial t_2}}_{HEE_2} < 0$, $\underbrace{\frac{\partial e_1^{PB}}{\partial t_2}}_{FEE_2} > 0$, $EDE_2 = (1-\gamma)D' + \underbrace{\frac{(1-\gamma)D'(FEE_2)}{\Lambda_2^{PB}}}_{(-)}$ and

$$\begin{aligned} \frac{dx_{21}^{PB}}{dt_2} &= \frac{2p_1' + p_1'' x_{11}}{J_1} < 0 \quad \frac{dx_{11}^{PB}}{dt_2} = -\frac{p_1' + p_1'' x_{11}}{J_1} > 0\,, \\ \frac{dx_{22}^{PB}}{dt_2} &= \frac{2p_2' + p_2'' x_{12}}{J_2} < 0 \quad \frac{dx_{12}^{PB}}{dt_2} = -\frac{p_2' + p_2'' x_{12}}{J_2} > 0\,. \end{aligned} \tag{13}$$

Under the PB-regime, both carbon taxes have the same negative impact on the consumer surplus in each country, $\partial CS_i/\partial t_i = \partial CS_i/\partial t_j = -p_i'(x_{1i} + x_{2i})\left(p_i'/J_i\right) < 0$. Hence, the CSE would call for a subsidy that increases the consumer surplus in both countries. However, this leads to positive consumer price spillovers (Haufler and Pflüger, 2007; Lockwood, 2001). That is, lowering taxes in one country benefits not only domestic but also foreign consumers. Since governments do not care about foreign consumers, this reduces the incentives of both governments to subsidise their consumers compared to the social optimum. Since governments choose their taxes non-cooperatively, focusing exclusively on the CSE, leads to a subsidy below the socially optimal level.

Under the non-cooperative regimes, the PSE reflects the concern of governments about the competitiveness of their firms. From (11) and (13), the production of each firm decreases (increases) in the tax of the home (foreign) country. As a result, the incentives related to producers are different from the cooperative solution. Setting a higher carbon tax reduces the sales of the home firm at the expenses of an expansion of the sales of the foreign firm. Hence, higher carbon taxes decrease the net profits (profits net of tax payments) of the home firm. In line with the analysis of Brander and Spencer (1985), each government has an incentive to give a subsidy to its firm in order to shift profits from the foreign to the home firm. This is known as the profit-shifting incentive. The PSE in both countries comprises two terms that reflect the incentive to shift profits in market 1 (first term) and market 2 (second term). Under this regime, both terms are negative, and, hence, would call for a subsidy.

Environmental damages and the problem of carbon leakage are captured by the EDE. Different from the social optimum, countries internalise only their individual damages under the non-cooperative regimes, i.e., a fraction of total damages. Both the home and the foreign PB-tax have the same effect on individual damages in both countries in that they reduce damages, i.e., $\partial D_i/\partial t_i = \partial D_i/\partial t_j = D_i' \left(p_1'/J_1 + p_2'/J_2\right) < 0$. If we exclusively focus on the EDE, ignoring other market distortions, the effectiveness of countries' climate policy can be measured by the departure of the EDE from individual marginal damages. The impact of a country's climate policy on global emissions comprises two effects: the "home emission effect" (HEE) and the "foreign emission effect" (FEE). HEE captures the change of emissions released by the home firm, $\partial e_i/\partial t_i$, and the FEE captures the change of emissions released by the foreign firm, $\partial e_j/\partial t_i$. The latter effect is sometimes referred to as the carbon leakage or transboundary pollution effect (Duval and Hamilton, 2002; Kennedy, 1994). Under the PB-regime, the HEE is negative but the FEE is positive, though overall the EDE is positive and hence the HEE is stronger than the FEE in absolute terms. Thus, carbon leakage weakens the EDE, which would call for a tax less than individual marginal damages, i.e., $EDE_i\,(PB) < D_i' \; \forall i = 1, 2$.[15] Recalling that the CSE and PSE are negative, it is clear that taxes under the PB-regime are set below individual marginal damages, $\hat{t}_i^{PB} < D_i' \; \forall\, i = 1, 2$.

In equilibrium, each government chooses its tax, given the tax level of the other government. That is, Nash equilibrium carbon taxes are obtained by solving $\partial W_1/\partial t_1 = 0$ and $\partial W_2/\partial t_2 = 0$ simultaneously. It is clear that since both countries face the same demand, the CSE is the same in both countries. If both firms were to produce the same quantities, the PSE would also be the same. Therefore, the difference in taxes is due to the EDE, which is higher in country 1 than country 2. Therefore, in equilibrium, country 1 will have a higher tax than country 2, i.e., $t_1^{PB*} > t_2^{PB*}$ for all $\gamma > 0.5$. As a result, equilibrium quantities produced by firm 1 are smaller than by firm 2 and, consequently, the producer surplus but also the net producer surplus (profits net of tax revenues) is smaller in country 1 than in country 2. Again, at this level of generality, nothing can be said whether equilibrium taxes will be positive or negative, except that, ceteris paribus, taxes in each country increase in the valuation of its individual marginal damages. Nevertheless, based on our discussion, we can state the following:

Proposition 1. *In a strategic trade model, the equilibrium production-based carbon tax will be lower than individual marginal damages in each country, with* $t_1^{PB*} \geq t_2^{PB*}$ *for all* $\gamma \geq 0.5$.

Given that both firms are assumed to be identical, different profits only stem from

[15] If pollution is local, the FEE will vanish and the EDE is equal to the country's marginal damage, see for instance Kennedy (1994).

differences in effective taxes that firms face. Hence, firm 1 suffers from a competitive disadvantage, both in the home and the foreign market. Therefore, in the subsequent analysis, we consider that country 1 complements its PB-policy with BCAs to protect its home firm and to raise the effectiveness of its climate policy. These regimes can be considered as a gradual shift to a unilateral consumption-based carbon tax imposed by country 1.

4.2 Border Carbon Adjustments on Imports (BI-regime)

Under the carbon adjustment of imports regime (BI-regime), country 1 complements its porduction-based tax with tariffs on imports. Tax rebates are not considered. From Table 1, the output levels in the second stage are given by $x_{11}^{BI} = x_{21}^{BI} = f(t_1, t_1) = g(t_1, t_1)$ for the supply to market 1, while $x_{12}^{BI} = f(t_1, t_2)$ and $x_{22}^{BI} = g(t_1, t_2)$ for the supply to market 2.

Maximising W_1^{BI} and W_2^{BI}, which are the welfare function of country 1 and country 2 under the BI-regime, with respect to own national taxes, leads to the optimal tax structure of each country (see Appendix A.4 for details):

$$\hat{t}_1^{BI} = \underbrace{\frac{p_1^{'}(x_{11}+x_{21})\frac{2p_1^{'}}{J_1}}{\Lambda_1^{BI}}}_{CSE_1\ (-)} + \underbrace{\frac{-p_1^{'}x_{11}\frac{dx_{21}^{BI}}{dt_1} - p_2^{'}x_{12}\frac{dx_{22}^{BI}}{dt_1}}{\Lambda_1^{BI}}}_{PSE_1 (+,-)} + \underbrace{\frac{\gamma\, D^{'}\left(\frac{\partial e^{BI}}{\partial t_1}\right)}{\Lambda_1^{BI}}}_{EDE_1\ (+)}$$
$$+ \underbrace{\frac{-x_{21} + t_2\frac{dx_{21}^{BI}}{dt_1}}{\Lambda_1^{BI}}}_{BCAIE_1 (+,?)}, \tag{14}$$

where $\Lambda_1^{BI} = \frac{dx_{11}^{BI}}{dt_1} + \frac{dx_{12}^{BI}}{dt_1} + \frac{dx_{21}^{BI}}{dt_1} < 0$, $\frac{\partial e^{BI}}{\partial t_1} = \frac{\partial e_1^{BI}}{\partial t_1} + \frac{\partial e_2^{BI}}{\partial t_1} < 0$, with $\underbrace{\frac{\partial e_1^{BI}}{\partial t_1}}_{HEE_1} < 0$, $\underbrace{\frac{\partial e_2^{BI}}{\partial t_1}}_{FEE_1} \gtreqless 0$, $EDE_1 = \gamma\, D^{'} + \underbrace{\frac{\gamma\, D^{'}\left(\frac{dx_{22}^{BI}}{dt_1}\right)}{\Lambda_1^{BI}}}_{(-)}$ and the effect of t_1 on outputs is given by:

$$\begin{aligned} &\frac{dx_{11}^{BI}}{dt_1} = \frac{dx_{21}^{BI}}{dt_1} = \frac{p_1^{'}}{J_1} < 0, \\ &\frac{dx_{12}^{BI}}{dt_1} = \frac{2p_2^{'} + p_2^{''}x_{22}}{J_2} < 0, \frac{dx_{22}^{BI}}{dt_1} = -\frac{p_2^{'} + p_2^{''}x_{22}}{J_2} > 0\,. \end{aligned} \tag{15}$$

$$\hat{t}_2^{BI} = \underbrace{\frac{p_2^{'}\left(x_{22}+x_{12}\right)\frac{p_2^{'}}{J_2}}{\Lambda_2^{BI}}}_{CSE_2\ (-)} + \underbrace{\frac{-p_2^{'}x_{22}\frac{dx_{12}^{BI}}{dt_2}}{\Lambda_2^{BI}}}_{PSE_2(-)} + \underbrace{\frac{(1-\gamma)\,D^{'}\left(\frac{\partial e^{BI}}{\partial t_2}\right)}{\Lambda_2^{BI}}}_{EDE_2\ (+)} + \underbrace{\frac{-x_{21}}{\Lambda_2^{BI}}}_{BCAIE_2(+)}\ , \tag{16}$$

where $\Lambda_2^{BI} = \frac{dx_{22}^{BI}}{dt_2} < 0$, $\frac{\partial e^{BI}}{\partial t_2} = \frac{\partial e_1^{BI}}{\partial t_2} + \frac{\partial e_2^{BI}}{\partial t_2} < 0$ with $\underbrace{\frac{\partial e_2^{BI}}{\partial t_2}}_{HEE_2} < 0$, $\underbrace{\frac{\partial e_1^{BI}}{\partial t_2}}_{FEE_2} > 0$,

$EDE_2 = (1-\gamma)D^{'} + \underbrace{\frac{(1-\gamma)D^{'}\left(\frac{dx_{12}^{BI}}{dt_2}\right)}{\Lambda_2^{BI}}}_{(-)}$ and the effect of t_2 on outputs is given by:

$$\begin{aligned} &\frac{dx_{21}^{BI}}{dt_2} = \frac{dx_{11}^{BI}}{dt_2} = 0\,, \\ &\frac{dx_{22}^{BI}}{dt_2} = \frac{2p_2^{'}+p_2^{''}x_{12}}{J_2} < 0 \ \frac{dx_{12}^{BI}}{dt_2} = -\frac{p_2^{'}+p_2^{''}x_{12}}{J_2} > 0\,. \end{aligned} \tag{17}$$

The CSE has the same sign as in the previous regime in both countries. However, t_1 has now a larger effect on the consumer surplus in country 1, while t_2 has no effect, $\partial CS_1/\partial t_1 = -p_1^{'}\left(x_{11}+x_{21}\right)\left(2p_1^{'}/J_1\right) < 0$ and $\partial CS_1/\partial t_2 = 0$. Therefore, country 1 has a stronger incentive to subsidise its consumers. For country 2, the consumer price spillovers to country 1 are zero, and, hence also country 2 has a stronger incentive to subsidise its consumers.

We now consider the PSE. We recall that one of the main objectives of BCAs is to protect profits in the light of asymmetric carbon taxes. Since the BI-regime assumes a unilateral BCA-policy imposed by country 1, we focus first on country 1, which, by assumption, sets a higher carbon tax than country 2. The sign of the PSE in country 1 becomes ambiguous after introducing carbon tariffs. In market 1, the profit-shifting incentive is eliminated and hence the first term of the PSE in (14) becomes now positive. BCAs on imports level the playing field in market 1 but the second term of the PSE is still negative, as firm 1 is not protected in market 2. Therefore, carbon tariffs reduce the pressure on country 1 of adjusting taxes downward. The sign of the aggregate PSE depends on the strength of these two effects which work in opposite directions. For a linear demand curve, the aggregate PSE is positive and hence would call for a tax.[16]

For country 2, its PSE is negative but comprises now only one component. This is because its carbon tax affects only the profits of its home firm in market 2, while

[16] See Appendix A.4 for details.

it has no effect on the profits obtained from exports to market 1 (as firm 2 de facto faces t_1 on its exports). Therefore, the profit-shifting incentive of country 2 in market 1 disappears and only that in market 2 remains.[17]

We now consider the EDE. With tariffs, the impact of country 1's tax t_1 on global emissions and carbon leakage becomes larger while the impact of country 2's tax t_2 becomes smaller. Whereas under the PB-regime, we had a symmetric impact of taxes on damages, $\partial D_i/\partial t_{i,j} = D_i^{'}\left(p_1^{'}/J_1 + p_2^{'}/J_2\right)$, now under the BI-regime we have $\partial D_i/\partial t_1 = D_i^{'}\left(2p_1^{'}/J_1 + p_2^{'}/J_2\right) > D_i^{'}\left(p_2^{'}/J_2\right) = \partial D_i/\partial t_2$. For both countries, the HEE remains negative. For country 2, also the FEE remains positive due to carbon leakage, however, the FEE in country 1 is now different. The FEE has two terms with opposite sign: $dx_{21}^{BI}/dt_1 < 0$ and $dx_{22}^{BI}/dt_1 > 0$. The first term may be viewed as anti-leakage, and, hence, irrespective of the sign of the overall FEE, carbon leakage can be better controlled by country 1. Hence, country 1 faces less pressure to reduce taxes to avoid leakage. Nevertheless, the EDE would still call for a tax lower than individual marginal damages in country 1. For country 2, an increase in its carbon tax raises only foreign emissions in market 2, while there is no effect on market 1. Hence, the FEE in country 2 stems only from $dx_{12}^{BI}/dt_2 > 0$. As a result, although carbon leakage is less sever under the BI-regime, also the EDE in country 2 would call for a tax below marginal damages. Hence, together, we have $EDE_i\left(BI\right) < D_i^{'}\ \forall i = 1,2$.

Finally, introducing BCAs on imports creates a new strategic incentive in both countries, which we call the BCAI-effect, abbreviated BCAIE. Recall that the effective carbon tariff rate depends on the difference between the two national taxes. Through the tariff, country 1 de facto taxes foreign production, which is a new source of revenues. This provides an incentive for country 1 to increase taxes. In contrast, country 2 has an incentive to tax exports of its firm in order to capture a larger part of its tax revenues. This constitutes a kind of a 'race to the top' in carbon taxes.

For country 1, the BCAIE comprises two terms. The first term, $-x_{21}/\Lambda_1^{BI}$, is strictly positive whereas the sign of second term, $\left(t_2 dx_{21}^{BI}/dt_1\right)/\Lambda_1^{BI}$, depends on the tax level of country 2. If $t_2 \geq 0$, the BCAIE in country 1 is unambiguously positive. However, if $t_2 < 0$, the sign of the second term is negative. The overall BCAIE could be negative but only if t_2 is negative and very small (i.e., large in absolute terms), implying a large subsidy. In such cases, if country 1 imposes a positive tax, the difference between the two national tax levels is large, implying

[17] It is not straightforward to confirm at this general level that the PSE in country 2 will strictly decrease because the absolute value of the denominator also decreases.

a large effective tariff rate, which may erode the BCA revenues in the sense of the Laffer curve. However, it is very likely that, overall, the BCAIE is positive, simply because if tax difference becomes too large, no interior equilibrium in output levels exists.

For country 2, the BCAIE is unambiguously positive and would call for taxing emissions. This is in line with the argument presented in Helm et al. (2012) who argue that the country facing a tariff reacts by imposing a carbon adjustment on exports.

Proposition 2. *In a strategic trade model, BCAs on imports create a new incentive for the country on which tariffs are levied on its exports to raise its tax. For the country which imposes tariffs, the pressure to lower taxes in order to shift profits and to countervail leakage effects is reduced.*

Although introducing BCAs changes some incentives towards a higher carbon tax in both countries, we cannot compare taxes across regimes at this level of generality. Also, comparing the two national taxes under this regime becomes ambiguous. As will be demonstrated later, even using specific functions, we have to impose a constraint such that $t_1^{BI*} > t_2^{BI*}$ in equilibrium.

Adding BCAs on imports eliminates the difference in profits of firms when supplying market 1. However, firm 1 still faces a competitive disadvantage in market 2. Therefore, we consider that country 1 will complement its BCA-policy on imports with a policy on exports.

4.3 Border Carbon Adjustments on Imports and Exports

Adding export rebates implies that firm 1 faces effective tax $t_1(1-\varphi)$ on its supply to market 2 with φ the rebate rate. We recall that WTO regulations require that $t_1(1-\varphi) \geq t_2$. We consider two forms of BCAs on exports. The BIE-regime assumes an optimal export adjustment/rebate parameter φ, which can be positive or negative and can be smaller or larger than 1, as explained in detail in Appendix A.5. The BF-regime assumes a full export rebate, i.e., $\varphi = 1$, and, hence, φ is not chosen optimally.

4.3.1 BCAs on Imports with Optimal Export Rebate (BIE-regime)

Under the BIE-regime, the output levels from the second stage are $x_{11}^{BIE} = x_{21}^{BIE} = f(t_1, t_1) = g(t_1, t_1)$ in market 1, and $x_{12}^{BIE} = f(t_1, t_2, \varphi)$ and $x_{22}^{BIE} = g(t_1, t_2, \varphi)$ in market 2 (see Table 1). Let W_1^{BIE} and W_2^{BIE} be the welfare function of country 1 and country 2, respectively, under this regime.

We derive the structure of the optimal export rebate rate by maximising W_1^{BIE} with respect to φ:

$$\hat{\varphi} = 1 + \underbrace{\frac{p_2^{'} x_{12} \frac{dx_{22}^{BIE}}{d\varphi}}{t_1 \frac{dx_{12}^{BIE}}{d\varphi}}}_{PSE_1(+)} + \underbrace{\frac{-\gamma\, D' \left(\frac{dx_{12}^{BIE}}{d\varphi} + \frac{dx_{22}^{BIE}}{d\varphi}\right)}{t_1 \frac{dx_{12}^{BIE}}{d\varphi}}}_{EDE_1(-)}, \tag{18}$$

where the effect of φ on market 2 is given by:

$$\begin{aligned} \frac{dx_{12}^{BIE}}{d\varphi} &= \frac{(-t_1)\,\left(2p_2^{'} + p_2^{''} x_{22}\right)}{J_2} > 0\,, \\ \frac{dx_{22}^{BIE}}{d\varphi} &= -\frac{(-t_1)\,\left(p_2^{'} + p_2^{''} x_{22}\right)}{J_2} < 0\,, \\ \frac{dX_2^{BIE}}{d\varphi} &= \frac{-t_1 p_2^{'}}{J_2} > 0\,. \end{aligned} \tag{19}$$

For simplicity, the signs in (18) and (19) as well as the subsequent interpretation assume positive values of t_1. See Appendix A.5 for further details and also for the case of negative values of t_1.[18]

The optimal export rebate rate depends on two opposing effects: the PSE, which calls for a large rebate rate to shift profits to firm 1 in market 2, and the EDE, which calls for a small rebate rate in order to reduce production and hence global emissions. From (18) it is evident that the optimal export rebate rate can be less than, equal or greater than a full rebate, i.e., $\hat{\varphi} \lesseqgtr 1$. The larger the EDE compared to the PSE, the smaller will be the optimal export rebate rate. A full rebate, i.e., $\hat{\varphi} = 1$, implying a unilateral consumption-based carbon tax, is only optimal if the PSE is equal to the EDE. Subsequently, we will first analyse the case of an optimal export rebate, ignoring the possibility of $\hat{\varphi} = 1$ for simplicity, as this case is covered under the next regime below. We recall that the optimal rebate rate can be larger than a full rebate if the climate policy level of country 2 is a subsidy.

Maximising W_1^{BIE} and W_2^{BIE} with respect to the national tax levels, delivers the

[18] In the example which we consider in Section 5, it will turn out that in equilibrium, $t_1 > 0$.

optimal tax structure of country 1 and country 2 (see details in Appendix A.5):

$$\hat{t}_1^{BIE} = \underbrace{\frac{p_1^{'}\left(x_{11}+x_{21}\right)\frac{2p_1^{'}}{J_1}}{\Lambda_1^{BIE}}}_{CSE_1\ (-)} + \underbrace{\frac{-p_1^{'}x_{11}\frac{dx_{21}^{BIE}}{dt_1} - p_2^{'}x_{12}\frac{dx_{22}^{BIE}}{dt_1}}{\Lambda_1^{BIE}}}_{PSE_1(+,?)} + \underbrace{\frac{\gamma\, D^{'}\left(\frac{\partial e^{BIE}}{\partial t_1}\right)}{\Lambda_1^{BIE}}}_{EDE_1\ (?)} + \underbrace{\frac{-x_{21}+t_2\frac{dx_{21}^{BIE}}{dt_1}}{\Lambda_1^{BIE}}}_{BCAIE_1\ (+,?)}, \tag{20}$$

where $\Lambda_1^{BIE} = \frac{dx_{11}^{BIE}}{dt_1} + \frac{dx_{21}^{BIE}}{dt_1} + (1-\varphi)\frac{dx_{12}^{BIE}}{dt_1} < 0$[19], $\frac{\partial e^{BIE}}{\partial t_1} \lesseqgtr 0$, with $\underbrace{\frac{\partial e_1^{BIE}}{\partial t_1}}_{HEE_1} \lesseqgtr 0$, $\underbrace{\frac{\partial e_2^{BIE}}{\partial t_1}}_{FEE_1} \gtreqless 0$, $EDE_1 = \gamma\, D^{'} + \underbrace{\frac{\gamma\, D^{'}\left(\varphi\frac{dx_{12}^{BIE}}{dt_1} + \frac{dx_{22}^{BIE}}{dt_1}\right)}{\Lambda_1^{BIE}}}_{(?)}$ and the effect of t_1 on outputs is:

$$\begin{aligned}
\frac{dx_{11}^{BIE}}{dt_1} &= \frac{dx_{21}^{BIE}}{dt_1} = \frac{p_1^{'}}{J_1} < 0,\\
\frac{dx_{12}^{BIE}}{dt_1} &= \frac{(1-\varphi)\left(2p_2^{'} + p_2^{''}x_{22}\right)}{J_2}\begin{cases} < 0 & if\, \varphi < 1 \\ > 0 & if\, \varphi > 1 \end{cases},\\
\frac{dx_{22}^{BIE}}{dt_1} &= -\frac{(1-\varphi)\left(p_2^{'} + p_2^{''}x_{22}\right)}{J_2}\begin{cases} > 0 & if\, \varphi < 1 \\ < 0 & if\, \varphi > 1 \end{cases}.
\end{aligned} \tag{21}$$

$$\hat{t}_2^{BIE} = \underbrace{\frac{p_2^{'}\left(x_{22}+x_{12}\right)\frac{p_2^{'}}{J_2}}{\Lambda_2^{BIE}}}_{CSE_2(-)} + \underbrace{\frac{-p_2^{'}x_{22}\frac{dx_{12}^{BIE}}{dt_2}}{\Lambda_2^{BIE}}}_{PSE_2(-)} + \underbrace{\frac{(1-\gamma)\, D^{'}\left(\frac{\partial e^{BIE}}{\partial t_2}\right)}{\Lambda_2^{BIE}}}_{EDE_2\ (+)} + \underbrace{\frac{-x_{21}}{\Lambda_2^{BIE}}}_{BCAIE_2(+)}, \tag{22}$$

where $\Lambda_2^{BIE} = \frac{dx_{22}^{BIE}}{dt_2} < 0$, $\frac{\partial e^{BIE}}{\partial t_2} = \frac{\partial e_1^{BIE}}{\partial t_2} + \frac{\partial e_2^{BIE}}{\partial t_2} < 0$, with $\underbrace{\frac{\partial e_2^{BIE}}{\partial t_2}}_{HEE_2} < 0$, $\underbrace{\frac{\partial e_1^{BIE}}{\partial t_2}}_{FEE_2} > 0$,

$EDE_2 = (1-\gamma)D^{'} + \underbrace{\frac{(1-\gamma)D^{'}\left(\frac{dx_{12}^{BIE}}{dt_2}\right)}{\Lambda_2^{BIE}}}_{(-)}$ and the effect of t_2 on outputs is similar

[19] We have: $\Lambda_1^{BIE} = \frac{2p_1^{'}}{J_1} + \frac{2p_2^{'}+p_2^{''}x_{22}}{J_2}\left(1-\varphi\right)^2 < 0$.

to the previous regime, namely:

$$\frac{dx_{21}^{BIE}}{dt_2} = \frac{dx_{11}^{BIE}}{dt_2} = 0\,,$$
$$\frac{dx_{22}^{BIE}}{dt_2} = \frac{2p_2^{'} + p_2^{''}x_{12}}{J_2} < 0 \quad \frac{dx_{12}^{BIE}}{dt_2} = -\frac{p_2^{'} + p_2^{''}x_{12}}{J_2} > 0\,. \tag{23}$$

Adding BCAs on exports affects the profit-shifting incentive of country 1 in market 2, which is the second term in the PSE in (20). On the one hand, if $\hat{\varphi} < 1$, the profit-shifting incentive in market 2 is negative, and, hence would call for a subsidy similar to the BI-regime. However, a higher t_1 will now induce a smaller increase in the market share of firm 2 in market 2, which, consequently, may reduce the incentive to shift profits by means of a low tax t_1. On the other hand, if $\hat{\varphi} > 1$, the second term in the PSE becomes positive, and, hence, would call for a tax. That is, the profit-shifting role works through exports rebates and firm 1 is overcompensated in its endeavour of competing in the foreign market.[20] In such cases, the overall PSE would be strictly positive, and, hence, would call for a higher tax than under the BI-regime.

The impact of complementing tariffs with an optimal export rebate is not straightforward on the EDE because export rebates raise the emissions of the home firm. We distinguish between two cases. First, if $\hat{\varphi} < 1$, a higher carbon tax in country 1 leads to a reduction in total emissions of the home firm. Hence, as usual, the HEE is negative. If the rebate rate is sufficiently large, the positive effect of t_1 on x_{22} is small so that it most likely to be offset by the negative effect of t_1 on x_{21}, such that the total emissions of firm 2 decrease and the FEE is also negative. Hence, we have $\partial e_1^{BIE}/\partial t_1 < 0$ and $\partial e_2^{BIE}/\partial t_1 < 0$, and the EDE could call for a tax larger than the individual marginal damages, i.e., $EDE_1 > D_1^{'}$ is possible. In other words, adding export rebates may lead to an overinternalisation of individual damages in country 1 in the absence of other effects. Second, if $\hat{\varphi} > 1$, a higher carbon tax in country 1 always reduces total emissions of the foreign country, i.e., the FEE is strictly negative. However, if $\hat{\varphi}$ is sufficiently large, the HEE could become positive. That is, total emissions released by firm 1 could increase in t_1 due to a high export rebate. At the extreme, global emissions could even increase in t_1 such that the EDE is negative. Obviously, these effects support the argument that adding export rebates may not address the environmental problem, at least if they are chosen too high.

[20] Overcompensating firms is not uncommon. For example, Martin et al. (2014) show that the free allocation of emission permits under the European Union Emissions Trading Scheme (EU-ETS) resulted in a sizeable overcompensation of emission-intensive industries.

The structure of the optimal climate policy of country 2 does not change after adding export rebates. However, as will be shown later, its equilibrium tax may well change due to the change of the reaction function of country 1.

4.3.2 BCAs on Imports with Full Export Rebate (BF-regime)

The only difference to the BIE-regime is that under the BF-regime the export rebate is set to $\varphi = 1$, which is equivalent to a unilateral consumption-based carbon policy. The output levels from the second stage are given by $x_{11}^{BF} = x_{21}^{BF} = f(t_1, t_1) = g(t_1, t_1)$ in market 1 and $x_{12}^{BF} = f(0, t_2)$ and $x_{22}^{BF} = g(0, t_2)$ in market 2.

Maximising the welfare functions of country 1 and 2, W_1^{BF} and W_2^{BF}, with respect to own national taxes gives the structure of optimal taxes (see Appendix A.6 for details):

$$\hat{t}_1^{BF} = \underbrace{\frac{p_1^{'}(x_{11}+x_{21})\frac{2p_1^{'}}{J_1}}{\Lambda_1^{BF}}}_{CSE_1\,(-)} + \underbrace{\frac{-p_1^{'}x_{11}\frac{dx_{21}^{BF}}{dt_1}}{\Lambda_1^{BF}}}_{PSE_1(+)} + \underbrace{\frac{\gamma\,D^{'}\left(\frac{\partial e^{BF}}{\partial t_1}\right)}{\Lambda_1^{BF}}}_{EDE_1\,(+)} + \underbrace{\frac{-x_{21}+t_2\frac{dx_{21}^{BF}}{dt_1}}{\Lambda_1^{BF}}}_{BCAIE_1\,(+,?)}$$
$$\Leftrightarrow -p_1^{''}x_{21}x_{11} + \frac{1}{2}t_2 + \gamma\,D^{'}\,, \tag{24}$$

where $\Lambda_1^{BF} = \frac{dx_{11}^{BF}}{dt_1} + \frac{dx_{21}^{BF}}{dt_1} < 0$, $\frac{\partial e^{BF}}{\partial t_1} < 0$, with $\underbrace{\frac{\partial e_1^{BF}}{\partial t_1}}_{HEE_1} = \underbrace{\frac{\partial e_2^{BF}}{\partial t_1}}_{FEE_1} < 0$, $EDE_1 = \gamma\,D^{'}$ and the effect of t_1 on both markets becomes:

$$\frac{dx_{11}^{BF}}{dt_1} = \frac{dx_{21}^{BF}}{dt_1} = \frac{p_1^{'}}{J_1} < 0\,,$$
$$\frac{dx_{12}^{BF}}{dt_1} = \frac{dx_{22}^{BF}}{dt_1} = 0\,. \tag{25}$$

$$\hat{t}_2^{BF} = \underbrace{\frac{p_2^{'}(x_{22}+x_{12})\frac{p_2^{'}}{J_2}}{\Lambda_2^{BF}}}_{CSE_2\,(-)} + \underbrace{\frac{-p_2^{'}x_{22}\frac{dx_{12}^{BF}}{dt_2}}{\Lambda_2^{BF}}}_{PSE_2\,(-)} + \underbrace{\frac{(1-\gamma)\;D^{'}\left(\frac{\partial e^{BF}}{\partial t_2}\right)}{\Lambda_2^{BF}}}_{EDE_2\,(+)} + \underbrace{\frac{-x_{21}}{\Lambda_2^{BF}}}_{BCAIE_2\,(+)}\,, \tag{26}$$

where $\Lambda_2^{BF} = \frac{dx_{22}^{BF}}{dt_2} < 0$, $\frac{\partial e^{BF}}{\partial t_2} < 0$, with $\underbrace{\frac{de_2^{BF}}{\partial t_2}}_{HEE_2} < 0$ and $\underbrace{\frac{de_1^{BF}}{\partial t_2}}_{FEE_2} > 0$, $EDE_2 = (1-\gamma)D^{'} + \underbrace{\frac{(1-\gamma)D^{'}\left(\frac{dx_{12}^{BF}}{dt_2}\right)}{\Lambda_2^{BF}}}_{(-)}$ and the effect of t_2 on outputs is similar to the previous two BCA-regimes:

$$\frac{dx_{21}^{BF}}{dt_2} = \frac{dx_{11}^{BF}}{dt_2} = 0\,,$$
$$\frac{dx_{22}^{BF}}{dt_2} = \frac{2p_2^{'} + p_2^{''}x_{12}}{J_2} < 0 \;\; \frac{dx_{12}^{BF}}{dt_2} = -\frac{p_2^{'} + p_2^{''}x_{12}}{J_2} > 0\,. \tag{27}$$

Under the BF-regime, the tax of each country targets only at the consumers in the home country. Hence, t_1 has no effect on the consumers in country 2, which raises the incentive of country 1 to subsidise its consumers, $\partial CS_2/\partial t_1 = 0$. The PSE in country 1 comprises now only one term, which would call for a tax to raise the net profits of the home firm obtained in market 1, while the effect of t_1 on market 2 vanishes. Therefore, the profit-shifting incentive has disappeared. In addition, country 1 completely controls the emissions released in the process of supplying its home market 1, without being offset by larger emissions released in the process of supplying market 2, either by its home or the foreign firm. Therefore, $\partial e_1^{BF}/\partial t_1 = \partial e_2^{BF}/\partial t_1 = p_1^{'}/J_1 < 0$. That is, the HEE and the FEE are equal and strictly negative. As a result, under this regime, the EDE in country 1 would call for a tax equal to its marginal damages, $EDE_1\,(BF) = D_1^{'}$.

Whether the optimal carbon tax of country 1 under this regime is equal to its individual marginal damages, depends on the demand curve and the tax level of country 2. It is clear from (24) that if the demand curve was linear ($p_1^{''} = 0$), and if country 2 was passive ($t_2 = 0$), full BCAs would restore the effectiveness of the carbon tax of country 1, where its optimal tax becomes equal to its individual marginal damages. However, if country 2 subsidises its production, $t_2 < 0$, also country will choose a tax below its marginal damage.

Proposition 3. *A production-based carbon tax supplemented by full BCAs is de facto a unilateral consumption-based tax, which corrects for both, the profit-shifting and carbon leakage distortions. The environmental damage effect calls for a tax equal to individual marginal damages in the environmentally more concerned country. The optimal tax level is positively correlated to the production-based carbon tax of the environmentally less concerned country.*

The optimal tax structure of country 2 is similar to the two previous BCA-regimes, implying that the structure of the optimal climate policy of country 2 is similar under the three BCA-regimes. However, the equilibrium tax level may nevertheless be different, depending on the reaction function of country 1 under the three BCA-regimes, as will become apparent in Section 5.

4.4 Bilateral Consumption-based Tax (CB-regime)

Finally, under the CB-regime, both countries impose a carbon tax on consumption, irrespective of the location of production. Therefore, the equilibrium output levels obtained from the second stage are $x_{11}^{CB} = x_{21}^{CB} = f(t_1, t_1) = g(t_1, t_1)$ for the supply of market 1 and $x_{12}^{CB} = x_{22}^{CB} = f(t_2, t_2) = g(t_2, t_2)$ for the supply of market 2. Let W_1^{CB} and W_2^{CB} be the welfare function of country 1 and country 2 under the CB-regime. (Details are provided in Appendix A.7.)

Maximising W_1^{CB} and W_2^{CB} with respect to own national taxes leads to following optimal tax structure:

$$\begin{aligned}
\hat{t}_1^{CB} &= \underbrace{\frac{p_1^{'}(x_{11}+x_{21})\frac{2p_1^{'}}{J_1}}{\Lambda_1^{CB}}}_{CSE_1\,(-)} + \underbrace{\frac{-p_1^{'}x_{11}\frac{dx_{21}^{CB}}{dt_1} - x_{21}}{\Lambda_1^{CB}}}_{PSE_1\,(+)} + \underbrace{\frac{\gamma\, D^{'}\left(\frac{\partial e^{CB}}{\partial t_1}\right)}{\Lambda_1^{CB}}}_{EDE_1\,(+)} \\
&\Leftrightarrow -\, p_1^{''}x_{21}x_{11} + \gamma\, D^{'}\,,
\end{aligned} \tag{28}$$

where $\Lambda_1^{CB} = \frac{dx_{11}^{CB}}{dt_1} + \frac{dx_{21}^{CB}}{dt_1} < 0,\ \frac{\partial e^{CB}}{\partial t_1} < 0$, with $\underbrace{\frac{\partial e_1^{CB}}{\partial t_1}}_{HEE_1} = \underbrace{\frac{\partial e_2^{CB}}{\partial t_1}}_{FEE_1} < 0,\ EDE_1 = \gamma\, D^{'}$ and the effect of t_1 on outputs is the same as under the BF-regime:

$$\begin{aligned}
\frac{\partial x_{11}^{CB}}{\partial t_1} &= \frac{\partial x_{21}^{CB}}{\partial t_1} = \frac{p_1^{'}}{J_1} < 0\,, \\
\frac{dx_{12}^{CB}}{dt_1} &= \frac{dx_{22}^{CB}}{dt_1} = 0\,.
\end{aligned} \tag{29}$$

$$\begin{aligned}
\hat{t}_2^{CB} &= \underbrace{\frac{p_2^{'}(x_{22}+x_{12})\frac{2p_2^{'}}{J_2}}{\Lambda_2^{CB}}}_{CSE_2\,(-)} + \underbrace{\frac{-p_2^{'}x_{22}\frac{dx_{12}^{CB}}{dt_2} - x_{12}}{\Lambda_2^{CB}}}_{PSE_2\,(+)} + \underbrace{\frac{(1-\gamma)\, D^{'}\left(\frac{\partial e^{CB}}{\partial t_2}\right)}{\Lambda_2^{CB}}}_{EDE_2\,(+)} \\
&\Leftrightarrow -p_2^{''}x_{12}x_{22} + (1-\gamma)\, D^{'}\,,
\end{aligned} \tag{30}$$

$\Lambda_2^{CB} = \frac{dx_{22}}{dt_2} + \frac{dx_{12}}{dt_2} < 0,\ \frac{\partial e^{CB}}{\partial t_2} < 0$, with $\underbrace{\frac{\partial e_2^{CB}}{\partial t_2}}_{HEE_2} = \underbrace{\frac{\partial e_1^{CB}}{\partial t_2}}_{FEE_2} < 0,\ EDE_2 = (1-\gamma)D^{'}$ and the effect of t_2 on market 2 changes, while on market 1 it remains as under the previous BCA-regimes.

$$\begin{aligned}
\frac{dx_{11}^{CB}}{dt_2} &= \frac{dx_{21}^{CB}}{dt_2} = 0\,, \\
\frac{\partial x_{22}^{CB}}{\partial t_2} &= \frac{\partial x_{12}^{CB}}{\partial t_2} = \frac{p_2^{'}}{J_2} < 0\,.
\end{aligned} \tag{31}$$

Under the CB-regime, there are no consumer price spillovers and the CSE would call for a subsidy which internalises the distortions stemming from underproduction efficiently (Haufler et al., 2005). That is, the CSE under the CB-regime and in the social optimum are the same (see (9), (28) and (30)).[21] In addition, the profit-shifting incentive in both countries is eliminated and the PSE is positive, calling to tax producers. The first term is the positive effect of a higher tax on the net profits of the home firm supplying the home market. The second term reflects the incentive of each country to shift tax revenues, recalling that the tax base under this regime is now $x_{1i} + x_{2i}$. Therefore, the competitiveness issue in both countries is totally solved under the CB-tax, where both firms share each market equally and profits of firms are equalised. Furthermore, the HEE and the FEE are also equalised implying that the climate policy in each country is effective to fully internalise own damages. That is, the EDE calls for consumption-based taxes which are equal to individual marginal damages, $EDE_i\,(CB) = D_i^{'}\;\forall\, i = 1, 2$.

The optimal level of a consumption-based tax in these models depends on the opposing incentives to care for consumers and producers. These in turn depend on the demand curvature as shown in Proposition 1 in Haufler et al. (2005). That is, if the demand curve is concave (convex), the incentive to tax producers (subsidise consumers) dominates, while the two incentives cancel out each other in the case of a linear demand curve.

Proposition 4. *While bilateral production-based carbon taxes are always set below individual marginal damages ($\hat{t}_i^{PB} < D_i^{'}$), bilateral consumption-based carbon taxes are set above (below) individual marginal damages if the demand curve is concave (convex): $\hat{t}_i^{CB} > (<)\, D_i^{'}$ if $p_i^{''} < (>)\, 0$, and equal to individual marginal damages if the demand curve is linear: $\hat{t}_i^{CB} = D_i^{'}$ if $p_i^{''} = 0$.*

Clearly, the difference between a unilateral consumption-based tax in (24) under the BF-regime and a bilateral consumption-based tax in (28) under the CB-regime for country 1 is due to the effect of the foreign climate policy level t_2, which are only identical if country 2 is passive, i.e., $\hat{t}_1^{BF} = \hat{t}_1^{CB}$ if $\hat{t}_2^{BF} = 0$.

To sum up, we showed in this section that a gradual shift from bilateral production-based carbon tax, along a unilateral consumption-based carbon tax to a bilateral consumption-based tax can partially or completely correct some distortions affecting the choice of carbon taxes. This concerns in particular the profit-shifting and carbon leakage effect. However, the incentives of governments

[21] The CSE in (9) is $2p_i^{'}x_{ki}$, and the CSE_1 in (28) is $2p_1^{'}x_{k1}$ given $x_{11} = x_{21}$. Similarly, the CSE_1 in (30) is $2p_2^{'}x_{k2}$ given $x_{12} = x_{22}$.

to subsidise consumers increases along this line. Hence, it is not possible to compare equilibrium carbon taxes at this level of generality. Therefore, in the next section, we use specific functions in order to rank equilibrium taxes across regimes.

5 Comparison of Equilibrium Climate Policies across Regimes

We assume a quadratic utility function $u_i(X_i) = aX_i - \frac{1}{2}X_i^2$, and, consequently, the inverse demand function for each country i is given by:

$$p_i = a - X_i\,, \quad \forall i = 1, 2\,, \tag{32}$$

where $a > 0$ is the chock-off price.

We also consider a strictly convex global damage function, which is given by:

$$D(e) = \frac{1}{2}de^2\,, \tag{33}$$

where $d > 0$ is a global damage parameter.

We provide all details of solving the two stages in Appendix B, including the derivation of the range of feasible parameter values for interior solutions under all regimes and the non-violation of WTO-rules under the three unilateral BCA-regimes, which are summarised in Table A.1.

The comparison of equilibrium carbon taxes depends only on two parameter values of the model; in particular, on the steepness of the global and individual marginal damage function related to parameter d, and the degree of asymmetry of damages among countries related to parameter γ.

Proposition 5. Ranking of Equilibrium Climate Policy Levels

1. *For country 1:*

   (a) $t_1^{PB*} < t_1^{BI*} < t_1^{BIE*}, t_1^{BF*}$ *with* $t_1^{BIE*} \leq (>)\, t_1^{BF*}$ *if* $\varphi^* \leq (>)\, 1$.

   (b) $t_1^{BIE*}, t_1^{BF*} < t_1^{CB*}$ *for all* $\gamma \geq \mathring{\gamma} = 0.73$.

2. *For country 2:*

   (a) $t_2^{BIE*}, t_2^{BF*} < t_2^{BI*} < t_2^{CB*}$ *with* $t_2^{BIE*} \geq (<)\, t_2^{BF*}$ *if* $\varphi^* \leq (>)\, 1$.[22]

[22] We have $t_2^{BIE*} = t_2^{BF*} = 0$ if $\gamma = 0.5$ irrespective of φ^*

(b) $t_2^{PB*} < t_2^{BIE*}, t_2^{BF*}$ *if either* $\gamma > \ddot{\gamma} \simeq 0.67$ *or if* $\gamma \leq \ddot{\gamma}$ *and the damage parameter* d *is sufficiently small, i.e.,* $d < \ddot{d}(\gamma)$.

3. $t_i^{PB*} < t_S^*$, *and* $t_i^{PB*} < t_i^{CB*}$ $\forall\, i = 1, 2$.

4. $t_S^* < t_i^{BI*}, t_i^{BIE*}, t_i^{BF*} < t_i^{CB*}$ $\forall\, i = 1, 2$ *if the damage parameter* d *is sufficiently low* $d < \tilde{d}(\gamma)$ *and countries are sufficiently asymmetric, i.e.,* $\gamma \geq \tilde{\gamma} \simeq 0.77$.

Proof. See Appendix B, including the precise definition of $\tilde{d}(\gamma)$ and $\ddot{d}(\gamma)$. □

The first result in (a) shows that the equilibrium carbon tax of country 1 increases gradually from the bilateral production-based tax to the unilateral BCA-regimes. If the equilibrium export rebate rate is less than a full rebate under the BIE-regime, i.e., $\varphi^* < 1$, the full export rebate under the BF-regime implies a larger carbon tax level in country 1, i.e., $t_1^{BIE*} < t_1^{BF*}$, and, vice-versa, if $\varphi^* > 1$, we have $t_1^{BIE*} > t_1^{BF*}$. As shown in Section 4, BCA measures reduce the rent shifting incentive of country 1, which reduces the pressure on country 1 to adjust its taxes downward. In addition, carbon leakage effect is mitigated, also reducing the pressure on country 1 to set low taxes. Furthermore, carbon tariffs are a new source of governmental income for country 1. Export rebates provide country 1 even with more control over carbon leakage, which explains that taxes under the BIE- and BF-regime are higher than under the BI-regime.

It is important to note that the ranking of taxes of country 1 are only the effective taxes firm 1 faces on its supply to country 1 under all regimes, i.e., $t_{11} = t_1$ as shown in Table 1. This is different for the effective tax firm 1 faces on its supply to country 2, t_{12}, under the BIE- and BF-regime, which depends on the rebate rate. This is immediately evident under the BF-regime with a full rebate, $\varphi = 1$, for which $t_{12}^{BF*} = 0$. However, also under the BIE-regime, we find $t_{12}^{BIE*} < t_1^{BIE*}$ as $t_1^{BIE*} > 0$ and $\varphi^* > 0$ (see Appendix B.4), and, by definition, we have $t_{12}^{BIE*} = t_1^{BIE*}(1 - \varphi^*)$.

It is also important to note that under the three BCA-regimes, $t_{21} = t_1$ (see Table 1) implying that effective carbon tax firm 2 faces on its supply to country 1 increases gradually under the three BCA-regimes. Therefore, by adding export rebates, country 1 shifts the cost of reducing emissions from the home to the foreign firm.

Whether in equilibrium the bilateral consumption-based tax (CB-regime) is larger than the partial or full unilateral consumption-based tax (BI-,BIE- and BF-regime) in country 1 depends on the parameters of the model, as shown in the first result in (b). If countries are sufficiently asymmetric, i.e., if $\gamma > \mathring{\gamma}$, the CB-tax of country 1 is larger than under the BCA-regimes. This is even the case for $\gamma \in [0.61, 0.73]$, provided d is sufficiently small.[23]

The second result shows the strategic role of BCAs on the carbon policy level of country 2. If countries are sufficiently asymmetric, i.e., $\gamma > \ddot{\gamma}$, country 2 sets a higher carbon tax under all BCA-regimes compared to the PB-regime. As mentioned above, under the BCA-regimes, the effective tax firm 2 faces on its supply to country 1 are those in country 1, $t_{21} = t_1$ (see Table 1). From the first result we have: $t_1^{BI*} < t_1^{BIE*}, t_1^{BF*}$ with $t_1^{BIE*} \leq (>)\, t_1^{BF*}$ if $\varphi^* \leq (>)\, 1$. In order to mitigate the negative effect of those taxes on firm 2, country 2 reacts by matching higher taxes t_1 with lower taxes t_2, which are the effective taxes its firm faces in its home market, i.e., $t_{22} = t_2$ (see Table 1). Hence, the ranking in the second result in (a) for the BCA-regimes is just the reverse of country 1. Thus, import tariffs only are better suited to induce country 2 to implement a higher carbon tax, whereas adding export rebates weakens this incentive. Therefore, our results show that adding export rebates not only raises emissions of firm 1 from exports, but also weakens the positive strategic effect induced by carbon tariffs on the equilibrium climate policy of country 2. Again, this suggests to consider export rebates cautiously. Interestingly, moving from a unilateral to a bilateral consumption-based tax implies higher taxes in country 2.

The last two results in Proposition 5 show that the PB-taxes are always set below the socially optimal tax and the CB-taxes in both countries. Although the incentive to subsidise consumers is lower under the PB-regime than under the CB-regime and in the social optimum, the absence of the profit-shifting effect and the full internalisation of individual (under the CB-regime) and global (in the social optimum) damages work in the opposite direction and are stronger. However, under the BCA- and the CB-regime, taxes in both countries may exceed the socially optimal tax level, though this is not generally the case. A sufficient condition for such a ranking as stated in Proposition 5 is that the individual and global marginal damage functions are not very steep (i.e., the value of parameter d is small) and countries are sufficiently asymmetric (i.e., the value of parameter γ is sufficiently large). In such cases, the EDE is less important in the social optimum compared to the CSE. Put differently, if governments behave non-cooperatively, the incentive to tax foreign production and/or to avoid carbon tariffs may lead to a 'race to the top' in non-cooperative taxes.

[23] Note that if we assumed a linear damage function, the CB-tax would always be larger than the taxes under the BCA-regimes in equilibrium.

6 Conclusions

Non-cooperative climate policies which regulate emissions by imposing a price of carbon through a production-based tax raises concerns about carbon leakage effects and the loss of competitiveness of home industries. In strategic trade models, it has been shown that these two concerns distort environmental policies to be inefficiently lax. In this paper, we analysed the effect of moving from a bilateral production-based carbon tax to a bilateral consumption-based tax, by considering in between partial and full unilateral consumption-based taxes, in the form of border carbon adjustments on imports (tariffs) and exports (rebates), on these distortions and equilibrium carbon taxes.

Apart from the social optimum, corresponding to fully cooperative regime, we considered five non-cooperative regimes. First, a bilateral production-based carbon tax, which is imposed by each government on its home firm (PB-regime). We then assumed that country 1, which gives a higher weight to environmental damages in our model, shifts gradually to a unilateral consumption-based tax using three forms of border carbon adjustments (BCAs), including carbon tariffs on imports (BI-regime) and carbon tariffs supplemented by export rebates, where the latter can be chosen optimally (BIE-regime) or chosen to be a full rebate (BF-regime). Under all three BCA-regimes, country 2 continues to impose its carbon tax on its home firm, choosing its tax optimally and strategically. Finally, we considered a bilateral consumption-based carbon tax (CB-regime). For each regime, we solved a two-stage game in which two countries first choose their carbon taxes, and then firms choose their equilibrium outputs, competing in both markets in a Nash-Cournot fashion.

We first derived the general optimal tax structure for both countries. We showed that BCA-measures, import tariffs and export rebates, could support more ambitious non-cooperative climate policies through three effects: 1) the profit-shifting incentives are reduced and/or eliminated, 2) carbon tariffs create a new incentive for both governments to tax emissions and 3) carbon leakage become less sever or even negative. We found that only BCAs on imports with a full export rebate could restore the effectiveness of country 1's carbon tax by fully internalising its own damages. However, there are two effects which point in the other direction. Import tariffs increase the incentive in both countries to subsidise their consumers as price spillover effects are reduced. Moreover, due to the strategic interaction among countries, government 2 responds to export rebates by reducing its tax and even choosing a subsidy level which, in equilibrium, could lead country 1 to set its tax below individual marginal damages. Country 2, which faces BCA-measures, always chooses its tax below its marginal damages. Only a bilateral consumption-based tax under the CB-regime implied that carbon taxes could be set equal or even above individual marginal damages.

We then ranked equilibrium carbon taxes across different regimes. For this, we had to assume specific functions. We found that the PB-taxes always fall short of the socially optimal tax, while both governments could impose non-cooperative carbon taxes under all other regimes including a bilateral consumption-based tax above those in the social optimum. However, this is only the case if countries are highly asymmetric in their perception of environmental damages and perceive environmental damages to be generally low. More important, if countries are sufficiently asymmetric, the three BCA-regimes implied that both countries set their carbon tax at a higher level than under the PB-regime. Although adding export rebates to import tariffs supports a higher carbon tax in country 1, though the reverse is true for country 2, which questions the effectiveness of export rebates for the internalisation of global damages.

Despite the fact that a bilateral consumption-based carbon tax eliminates both the concern about a loss of competitiveness and carbon leakage, an agreement is needed among countries to switch to this regime. However, this might face a coordination problem, similar to switching from a non-cooperative to cooperative carbon tax regime, which may face objections by the environmentally less concerned government. Therefore, it might be expected that a shift to unilateral consumption-based taxes through BCA-measures is more likely to come about in the future. Furthermore, we showed that the strategic role of BCAs to support a stricter non-cooperative policy level in country 2 is triggered by carbon tariffs, whereas export rebates weaken this role.

Several extensions could be considered in future work. First, it might be interesting to examine under which conditions the CB-regime constitutes a Pareto-improvement to both governments compared to the PB-regime. Second, we considered in this paper the effect of different forms of BCAs on non-cooperative carbon taxes. However, one could also examine the effect of these measures on enforcing cooperation among countries. Third, we assume the location of firms to be fixed. Hence, we considered the effect of BCAs on one channel of carbon leakage which is the relocation of production through trade. However, firms may completely close down and relocate their production facilities abroad to countries with less strict climate policies. Therefore, a possible extension is to allow for the endogenous choice of the location of firms along the lines proposed by Markusen et al. (1995) and Hoel (1997).

References

Babiker, M. H. and Rutherford, T. F. (2005). The economic effects of border measures in subglobal climate agreements. *The Energy Journal*, 26(4):99–125.

Barnett, A. H. (1980). The Pigouvian tax rule under monopoly. *The American Economic Review*, 70(5):1037–1041.

Barrett, S. (1994). Strategic environmental policy and intrenational trade. *Journal of Public Economics*, 54(3):325–338.

Böhringer, C., Balistreri, E. J., and Rutherford, T. F. (2012). The role of border carbon adjustment in unilateral climate policy: overview of an energy modeling forum study EMF 29. *Energy Economics*, 34:S97–S110.

Böhringer, C., Fischer, C., and Rosendahl, K. E. (2014). Cost-effective unilateral climate policy design: size matters. *Journal of Environmental Economics and Management*, 67(3):318–339.

Brander, J. A. and Spencer, B. J. (1985). Export subsidies and international market share rivalry. *Journal of International Economics*, 18(1-2):83–100.

Branger, F. and Quirion, P. (2014). Would border carbon adjustments prevent carbon leakage and heavy industry competitiveness losses? Insights from a meta-analysis of recent economic studies. *Ecological Economics*, 99:29–39.

Conrad, K. (1993). Taxes and subsidies for pollution-intensive industries as trade policy. *Journal of Environmental Economics and Management*, 25(2):121–135.

Copeland, B. R. (1996). Pollution content tariffs, environmental rent shifting, and the control of cross-border pollution. *Journal of International Economics*, 40(3-4):459–476.

Duval, Y. and Hamilton, S. F. (2002). Strategic environmental policy and international trade in asymmetric oligopoly markets. *International Tax and Public Finance*, 9(3):259–271.

Eichberger, J. (1993). *Game Theory for Economists*. Academic Press.

Eichner, T. and Pethig, R. (2015). Unilateral consumption-based carbon taxes and negative leakage. *Resource and Energy Economics*, 40:127–142.

Eyland, T. and Zaccour, G. (2012). Strategic effects of a border tax adjustment. *International Game Theory Review*, 14(03):1250016.

Eyland, T. and Zaccour, G. (2014). Carbon tariffs and cooperative outcomes. *Energy Policy*, 65:718–728.

Fischer, C. and Fox, A. K. (2012). Comparing policies to combat emissions leakage: border carbon adjustments versus rebates. *Journal of Environmental Economics and Management*, 64(2):199–216.

Friedman, J. W. (1986). *Game Theory with Applications to Economics*. Oxford University Press, USA.

Haufler, A. and Pflüger, M. (2007). International oligopoly and the taxation of commerce with revenue-constrained governments. *Economica*, 74(295):451–473.

Haufler, A., Schjelderup, G., and Stähler, F. (2005). Barriers to trade and imperfect competition: the choice of commodity tax base. *International Tax and Public Finance*, 12(3):281–300.

Hecht, M. and Peters, W. (2018). Border adjustments supplementing nationally determined carbon pricing. *Environmental and Resource Economics*, 73(1):93–109.

Helm, D., Hepburn, C., and Ruta, G. (2012). Trade, climate change, and the political game theory of border carbon adjustments. *Oxford Review of Economic Policy*, 28(2):368–394.

Hoel, M. (1996). Should a carbon tax be differentiated across sectors? *Journal of Public Economics*, 59(1):17–32.

Hoel, M. (1997). Environmental policy with endogenous plant locations. *The Scandinavian Journal of Economics*, 99(2):241–259.

Jakob, M., Marschinski, R., and Hübler, M. (2013). Between a rock and a hard place: a trade-theory analysis of leakage under production-and consumption-based policies. *Environmental and Resource Economics*, 56(1):47–72.

Jakob, M., Steckel, J. C., and Edenhofer, O. (2014). Consumption-versus production-based emission policies. *Annual Review of Resource Economics*, 6(1):297–318.

Kennedy, P. W. (1994). Equilibrium pollution taxes in open economies with imperfect competition. *Journal of Environmental Economics and Management*, 27(1):49–63.

Lockwood, B. (2001). Tax competition and tax co-ordination under destination and origin principles: a synthesis. *Journal of Public Economics*, 81(2):279–319.

Markusen, J. R. (1975). International externalities and optimal tax structures. *Journal of International Economics*, 5(1):15–29.

Markusen, J. R., Morey, E. R., and Olewiler, N. (1995). Competition in regional environmental policies when plant locations are endogenous. *Journal of Public Economics*, 56(1):55–77.

Martin, R., Muûls, M., De Preux, L. B., and Wagner, U. J. (2014). Industry compensation under relocation risk: a firm-level analysis of the EU emissions trading scheme. *American Economic Review*, 104(8):2482–2508.

Nicolaï, J.-P., Péchoux, I., Ponssard, J.-P., and Pouyet, J. (2010). Environmental policy and border adjustments with imperfect competition. *Revue Économique*, 61(1):57–77.

Peters, G. P. and Hertwich, E. G. (2008). Post-Kyoto greenhouse gas inventories: production versus consumption. *Climatic Change*, 86(1-2):51–66.

Sanctuary, M. (2018). Border carbon adjustments and unilateral incentives to regulate the climate. *Review of International Economics*, 26(4):826–851.

Steininger, K., Lininger, C., Droege, S., Roser, D., Tomlinson, L., and Meyer, L. (2014). Justice and cost effectiveness of consumption-based versus production-based approaches in the case of unilateral climate policies. *Global Environmental Change*, 24:75–87.

Stiglitz, J. (2006). A new agenda for global warming. *The Economists' Voice*, 3(7).

Weitzel, M., Hübler, M., and Peterson, S. (2012). Fair, optimal or detrimental? Environmental vs. strategic use of border carbon adjustment. *Energy Economics*, 34:S198–S207.

Yomogida, M. and Tarui, N. (2013). Emission taxes and border tax adjustments for oligopolistic industries. *Pacific Economic Review*, 18(5):644–673.

Appendix

A General Functions

A.1 The Second Stage

The second-order conditions of profit maximisation in (4) and (5) are given by:

$\frac{\partial^2 \pi_{11}}{\partial x_{11}^2} = 2p_1^{'} + x_{11}p_1^{''} < 0$ and $\frac{\partial^2 \pi_{21}}{\partial x_{21}^2} = 2p_1^{'} + x_{21}p_1^{''} < 0$, as well as

$\frac{\partial^2 \pi_{12}}{\partial x_{12}^2} = 2p_2^{'} + x_{12}p_2^{''} < 0$ and $\frac{\partial^2 \pi_{22}}{\partial x_{22}^2} = 2p_2^{'} + x_{22}p_2^{''} < 0$,

which we assume to hold. Together with the condition imposed in the paper for goods to be strategic substitutes, this implies:

$$J_1 = \begin{bmatrix} \frac{\partial^2 \pi_{11}}{\partial x_{11}^2} & \frac{\partial^2 \pi_{11}}{\partial x_{11} \partial x_{21}} \\ \frac{\partial^2 \pi_{21}}{\partial x_{21} \partial x_{11}} & \frac{\partial^2 \pi_{21}}{\partial x_{21}^2} \end{bmatrix} = p_1^{'}(3p_1^{'} + p_1^{''}(x_{11} + x_{21})) > 0 \text{ for market 1 and}$$

$$J_2 = \begin{bmatrix} \frac{\partial^2 \pi_{12}}{\partial x_{12}^2} & \frac{\partial^2 \pi_{12}}{\partial x_{12} \partial x_{22}} \\ \frac{\partial^2 \pi_{22}}{\partial x_{22} \partial x_{12}} & \frac{\partial^2 \pi_{22}}{\partial x_{22}^2} \end{bmatrix} = p_2^{'}(3p_2^{'} + p_2^{''}(x_{12} + x_{22})) > 0 \text{ for market 2.}$$

The above conditions ensure the existence and uniqueness of a Nash equilibrium (Eichberger, 1993; Friedman, 1986) in the second stage. If those conditions hold globally, they are also sufficient for the Routh-Hurwitz stability condition to be satisfied (Brander and Spencer, 1985).

A.2 Social Optimum

From the second stage, both firms face a uniform tax to supply each market, i.e., $x_{1i}^S = x_{2i}^S = f(t_S, t_S) = g(t_S, t_S)$.

The first-order conditions of profit-maximisation in (4) for market 1 can be written as:

$\frac{\partial \pi_{11}}{\partial x_{11}} = p_1 + x_{11}p_1^{'} - c - t_S = 0$ & $\frac{\partial \pi_{21}}{\partial x_{21}} = p_1 + x_{21}p_1^{'} - c - t_S = 0$.

In order to derive the effect of t_S on equilibrium output of firms, we totally differentiate the first-order conditions:

$\frac{\partial^2 \pi_{11}}{\partial x_{11}^2} dx_{11} + \frac{\partial^2 \pi_{11}}{\partial x_{11}\partial x_{21}} dx_{21} + \frac{\partial^2 \pi_{11}}{\partial x_{11}\partial t_S} dt_S = 0$, and

$\frac{\partial^2 \pi_{21}}{\partial x_{21}\partial x_{11}} dx_{11} + \frac{\partial^2 \pi_{21}}{\partial x_{21}^2} dx_{21} + \frac{\partial^2 \pi_{21}}{\partial x_{21}\partial t_S} dt_S = 0$.

From Appendix A.1, and by substituting $\frac{\partial^2 \pi_{11}}{\partial x_{11}\partial t_S} = \frac{\partial^2 \pi_{21}}{\partial x_{21}\partial t_S} = -1$, we obtain:

$\begin{bmatrix} 2p_1^{'} + x_{11}p_1^{''} & p_1^{'} + x_{11}p_1^{''} \\ p_1^{'} + x_{21}p_1^{''} & 2p_1^{'} + x_{21}p_1^{''} \end{bmatrix} \begin{bmatrix} dx_{11} \\ dx_{21} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} dt_S$.

By applying Cramer's rule, the effect of t_S on x_{11} and x_{21} is given by:

$\frac{dx_{11}}{dt_S} = \begin{bmatrix} 1 & p_1^{'} + x_{11}p_1^{''} \\ 1 & 2p_1^{'} + x_{21}p_i^{''} \end{bmatrix} / J_1 = \frac{p_1^{'}}{J_1} < 0$ given $x_{11} = x_{21}$ and $J_1 > 0$,

$\frac{dx_{21}}{dt_S} = \begin{bmatrix} 2p_1^{'} + x_{11}p_1^{''} & 1 \\ p_1^{'} + x_{21}p_1^{''} & 1 \end{bmatrix} / J_1 = \frac{p_1^{'}}{J_1} < 0$ given $x_{21} = x_{11}$ and $J_1 > 0$.

Following the same steps for market 2, we obtain $\frac{dx_{12}}{dt_S} = \frac{dx_{22}}{dt_S} = \frac{p_2^{'}}{J_2} < 0$.

The aggregate welfare function is given by:

$$\begin{aligned} W = {} & u_1(X_1) - p_1X_1 + u_2(X_2) - p_2X_2 + \Pi_1 + \Pi_2 \\ & + t_S(x_{11} + x_{12}) + t_S(x_{22} + x_{21}) - D_1(e) - D_2(e) . \end{aligned} \tag{A.1}$$

Differentiation of (A.1) with respect to t_S gives:

$\frac{\partial CS_i}{\partial t_S} = u_i^{'}\left(\frac{dx_{1i}}{dt_S} + \frac{dx_{2i}}{dt_S}\right) - \left[p_i\left(\frac{dx_{1i}}{dt_S} + \frac{dx_{2i}}{dt_S}\right) + \left(p_i^{'}\left(\frac{dx_{1i}}{dt_S} + \frac{dx_{2i}}{dt_S}\right)(x_{1i} + x_{2i})\right)\right]$.

From (1) in the text, we have $p_i = u_i^{'}$, and therefore:

$\frac{\partial CS_i}{\partial t_S} = -p_i^{'}\left(\frac{dx_{1i}}{dt_S} + \frac{dx_{2i}}{dt_S}\right)(x_{1i} + x_{2i}) < 0\,$, and

$\frac{\partial \Pi_k}{\partial t_S} = \frac{\partial \pi_{k1}}{\partial t_S} + \frac{\partial \pi_{k2}}{\partial t_S}$

$\quad = (p_1 - c - t_S)\frac{dx_{k1}}{dt_S} + x_{k1}\left[p_1^{'}\left(\frac{dx_{k1}}{dt_S} + \frac{dx_{\ell 1}}{dt_S}\right) - 1\right]$

$\quad + (p_2 - c - t_S)\frac{dx_{k2}}{dt_S} + x_{k2}\left[p_2^{'}\left(\frac{dx_{k2}}{dt_S} + \frac{dx_{\ell 2}}{dt_S}\right) - 1\right].$

The national carbon tax has three effects on the profits of the home firm: a change in production, a change in prices and a change of tax payments. The effect of the tax on prices is through the domestic sales and imports. In the cooperative solution, an increase in the socially optimal tax reduces both, the domestic and the foreign production. From (4) and (5), $(p_1 - c - t_S) = -x_{k1}p_1^{'}$ and, similarly, $(p_2 - c - t_S) = -x_{k2}p_2^{'}$. A reduction in domestic production and an increase in prices through the domestic sales cancels out, $(p_i - c - t_S)\frac{\partial x_{ki}}{\partial t_S} + x_{ki}p_i^{'}\frac{\partial x_{ki}}{\partial t_S} = 0$. Therefore, the net effect is an increase in the market price due to a reduction of imports, $x_{ki}p_i^{'}\frac{\partial x_{\ell i}}{\partial t_S} > 0$, and the tax payments x_{ki}.

Therefore, $\frac{\partial \Pi_k}{\partial t_S} = p_1^{'}x_{k1}\frac{dx_{\ell 1}}{dt_S} + p_2^{'}x_{k2}\frac{dx_{\ell 2}}{dt_S} - x_{k1} - x_{k2}$,

$\frac{\partial TR_i}{\partial t_S} = t_S\left(\frac{dx_{k1}}{dt_S} + \frac{dx_{k2}}{dt_S}\right) + x_{k1} + x_{k2}$, and

$\frac{\partial D_1}{t_S} + \frac{\partial D_2}{t_S} = \gamma D^{'}\left(\frac{\partial e}{\partial X}\frac{dX}{dt_S}\right) + (1-\gamma)\,D^{'}\left(\frac{\partial e}{\partial X}\frac{dX}{dt_S}\right) = D^{'}\frac{dX}{dt_S}$, where $e = X = X_1 + X_2$ and $\frac{\partial e}{\partial X} = 1$, given that we normalise the emission output coefficient to 1.

Since both firms produce the same amount, each firm divides its production for both markets equally and since we assume symmetric utility functions, $\frac{dX}{dt_S} = \frac{dX_1}{dt_S} + \frac{dX_2}{dt_S} < 0$.

Therefore, the first-order condition derived from maximising (A.1) with respect to t_S can be written as:

$$\frac{\partial W}{\partial t_S} = -2p_i^{'}\,x_{ki}\left(\frac{dX}{dt_S}\right) + p_i^{'}x_{ki}\left(\frac{dX}{dt_S}\right) - D^{'}\left(\frac{dX}{dt_S}\right) + t_S\left(\frac{dX}{dt_S}\right) = 0\,, \qquad \text{(A.2)}$$

or, simplifying: $\frac{\partial W}{\partial t_S} = -p_i^{'}x_{ki}\left(\frac{dX}{dt_S}\right) + \left(t_S - D^{'}\right)\left(\frac{dX}{dt_S}\right) = 0$, which leads to equation (9) in the text.

A.3 PB-regime

The first-order conditions of profit-maximisation in (4) and (5) can be written as:

$\frac{\partial \pi_{11}}{\partial x_{11}} = p_1 + x_{11}p_1^{'} - c - t_1 = 0$ & $\frac{\partial \pi_{21}}{\partial x_{21}} = p_1 + x_{21}p_1^{'} - c - t_2 = 0\,,$
$\frac{\partial \pi_{12}}{\partial x_{12}} = p_2 + x_{12}p_2^{'} - c - t_1 = 0$ & $\frac{\partial \pi_{22}}{\partial x_{22}} = p_2 + x_{22}p_2^{'} - c - t_2 = 0\;.$

The effect of t_1 on production levels are given by total differentiation of the above functions. In market 1, we have:

$\frac{\partial^2\pi_{11}}{\partial x_{11}^2}dx_{11}+\frac{\partial^2\pi_{11}}{\partial x_{11}\partial x_{21}}dx_{21}+\frac{\partial^2\pi_{11}}{\partial x_{11}\partial t_1}dt_1=0$, and

$\frac{\partial^2\pi_{21}}{\partial x_{21}\partial x_{11}}dx_{11}+\frac{\partial^2\pi_{21}}{\partial x_{21}^2}dx_{21}+\frac{\partial^2\pi_{21}}{\partial x_{21}\partial t_1}dt_1=0$.

From Appendix A.1 and by substituting $\frac{\partial^2\pi_{11}}{\partial x_{11}\partial t_1}=-1$ and $\frac{\partial^2\pi_{21}}{\partial x_{21}\partial t_1}=0$, we obtain:

$\left[\begin{array}{cc} 2p_1^{'}+x_{11}p_1^{''} & p_1^{'}+x_{11}p_1^{''} \\ p_1^{'}+x_{21}p_1^{''} & 2p_1^{'}+x_{21}p_1^{''}\end{array}\right]\left[\begin{array}{c} dx_{11} \\ dx_{21}\end{array}\right]=\left[\begin{array}{c} 1 \\ 0\end{array}\right]dt_1$,

$\frac{dx_{11}^{PB}}{dt_1}=\left[\begin{array}{cc} 1 & p_1^{'}+x_{11}p_1^{''} \\ 0 & 2p_1^{'}+x_{21}p_1^{''}\end{array}\right]/J_1=\frac{2p_1^{'}+x_{21}p_1^{''}}{J_1}<0$,

$\frac{dx_{21}^{PB}}{dt_1}=\left[\begin{array}{cc} 2p_1^{'}+x_{11}p_1^{''} & 1 \\ p_1^{'}+x_{21}p_1^{''} & 0\end{array}\right]/J_1=-\frac{p_1^{'}+x_{21}p_1^{''}}{J_1}>0$, and $\frac{dx_{11}^{PB}}{dt_1}+\frac{dx_{21}^{PB}}{dt_1}=\frac{p_1^{'}}{J_1}$, where $J_1>0$.

In market 2, we have:

$\frac{\partial^2\pi_{12}}{\partial x_{12}^2}dx_{12}+\frac{\partial^2\pi_{12}}{\partial x_{12}\partial x_{22}}dx_{22}+\frac{\partial^2\pi_{12}}{\partial x_{12}\partial t_1}dt_1=0$, and

$\frac{\partial^2\pi_{22}}{\partial x_{22}\partial x_{12}}dx_{12}+\frac{\partial^2\pi_{22}}{\partial x_{22}^2}dx_{22}+\frac{\partial^2\pi_{22}}{\partial x_{22}\partial t_1}dt_1=0$.

From Appendix A.1, and by substituting $\frac{\partial^2\pi_{12}}{\partial x_{12}\partial t_1}=-1$ and $\frac{\partial^2\pi_{22}}{\partial x_{22}\partial t_1}=0$, we obtain:

$\left[\begin{array}{cc} 2p_2^{'}+x_{12}p_2^{''} & p_2^{'}+x_{12}p_2^{''} \\ p_2^{'}+x_{22}p_2^{''} & 2p_2^{'}+x_{22}p_2^{''}\end{array}\right]\left[\begin{array}{c} dx_{12} \\ dx_{22}\end{array}\right]=\left[\begin{array}{c} 1 \\ 0\end{array}\right]dt_1$,

$\frac{dx_{12}^{PB}}{dt_1}=\left[\begin{array}{cc} 1 & p_2^{'}+x_{12}p_2^{''} \\ 0 & 2p_2^{'}+x_{22}p_2^{''}\end{array}\right]/J_2=\frac{2p_2^{'}+x_{22}p_2^{''}}{J_2}<0$,

$\frac{dx_{22}^{PB}}{dt_1}=\left[\begin{array}{cc} 2p_2^{'}+x_{12}p_2^{''} & 1 \\ p_2^{'}+x_{22}p_2^{''} & 0\end{array}\right]/J_2=-\frac{p_2^{'}+x_{22}p_2^{''}}{J_2}>0$, and $\frac{dx_{12}^{PB}}{dt_1}+\frac{dx_{22}^{PB}}{dt_1}=\frac{p_2^{'}}{J_2}$, where $J_2>0$.

These are the effects of t_1 on both markets as given in (11) in the text. The effect of t_2 on market 1 and 2 in (13) is obtained in a similar way.

The welfare function of each country is given by :

$$W_i^{PB}=u_i\left(X_i\right)-p_iX_i+\Pi_k+t_i\left(x_{k1}+x_{k2}\right)-D_i\left(e\right). \tag{A.3}$$

Maximising W_1^{PB} with respect to t_1 leads to the following first-order condition:

$$\begin{aligned}\frac{\partial W_1^{PB}}{\partial t_1} &= -p_1^{'}\left(\frac{dx_{11}^{PB}}{dt_1}+\frac{dx_{21}^{PB}}{dt_1}\right)(x_{11}+x_{21})+p_1^{'}x_{11}\frac{dx_{21}^{PB}}{dt_1}+p_2^{'}x_{12}\frac{dx_{22}^{PB}}{dt_1}\\ &\quad +t_1\left(\frac{dx_{11}^{PB}}{dt_1}+\frac{dx_{12}^{PB}}{dt_1}\right)-\gamma D^{'}\left(\frac{dx_{11}^{PB}}{dt_1}+\frac{dx_{12}^{PB}}{dt_1}+\frac{dx_{22}^{PB}}{dt_1}+\frac{dx_{21}^{PB}}{dt_1}\right)=0\,,\end{aligned} \tag{A.4}$$

which can be written as:

$-p_1^{'}\left(\frac{dx_{11}^{PB}}{dt_1}+\frac{dx_{21}^{PB}}{dt_1}\right)(x_{11}+x_{21})+p_1^{'}x_{11}\frac{dx_{21}^{PB}}{dt_1}+p_2^{'}x_{12}+\left(t_1-\gamma D^{'}\right)\left(\frac{dx_{11}^{PB}}{dt_1}+\frac{dx_{12}^{PB}}{dt_1}\right)-\gamma D^{'}\left(\frac{dx_{22}^{PB}}{dt_1}+\frac{dx_{21}^{PB}}{dt_1}\right)=0\,,$

where,

$\frac{\partial CS_1}{\partial t_1}=-p_1^{'}\left(\frac{dx_{11}^{PB}}{dt_1}+\frac{dx_{21}^{PB}}{dt_1}\right)(x_{11}+x_{21})<0,$

$\frac{\partial \Pi_1}{\partial t_1}=p_1^{'}x_{11}\frac{dx_{21}^{PB}}{dt_1}+p_2^{'}x_{12}\frac{dx_{22}^{PB}}{dt_1}-x_{11}-x_{12}<0,$

$\frac{\partial TR_1}{\partial t_1}=t_1\left(\frac{dx_{11}^{PB}}{dt_1}+\frac{dx_{12}^{PB}}{dt_1}\right)+x_{11}+x_{12}$, and

$\frac{\partial D_1}{\partial t_1}=\gamma D^{'}\left(\frac{\partial e_1}{\partial t_1}+\frac{\partial e_2}{\partial t_2}\right)=\gamma D^{'}\left(\frac{\partial e_1}{\partial x_{11}}\frac{dx_{11}^{PB}}{dt_1}+\frac{\partial e_1}{\partial x_{12}}\frac{dx_{12}^{PB}}{dt_1}+\frac{\partial e_2}{\partial x_{22}}\frac{dx_{22}^{PB}}{dt_1}+\frac{\partial e_2}{\partial x_{21}}\frac{dx_{21}^{PB}}{dt_1}\right)=\gamma D^{'}\left(\frac{dx_{11}^{PB}}{dt_1}+\frac{dx_{12}^{PB}}{dt_1}+\frac{dx_{22}^{PB}}{dt_1}+\frac{dx_{21}^{PB}}{dt_1}\right)<0\,.$

Having $\left(t_1-\gamma D^{'}\right)\left(\frac{dx_{11}^{PB}}{dt_1}+\frac{dx_{12}^{PB}}{dt_1}\right)-\gamma D^{'}\left(\frac{dx_{22}^{PB}}{dt_1}+\frac{dx_{21}^{PB}}{dt_1}\right)$ as stated above, the EDE_1 in (10) can be written as $\gamma\,D^{'}+\frac{\gamma\,D^{'}\left(\frac{dx_{22}^{PB}}{dt_1}+\frac{dx_{21}^{PB}}{dt_1}\right)}{\Lambda_1^{PB}}$.

Similarly, for country 2, we obtain:

$$\begin{aligned}\frac{\partial W_2^{PB}}{\partial t_2} &= -p_2^{'}\left(\frac{dx_{22}^{PB}}{dt_2}+\frac{dx_{12}^{PB}}{dt_2}\right)(x_{22}+x_{12})+p_2^{'}x_{22}\frac{dx_{12}^{PB}}{dt_2}+p_1^{'}x_{21}\frac{dx_{11}^{PB}}{dt_2}\\ &\quad +t_2\left(\frac{dx_{22}^{PB}}{dt_2}+\frac{dx_{21}^{PB}}{dt_2}\right)-(1-\gamma)\,D^{'}\left(\frac{dx_{22}^{PB}}{dt_2}+\frac{dx_{21}^{PB}}{dt_2}+\frac{dx_{11}^{PB}}{dt_2}+\frac{dx_{12}^{PB}}{dt_2}\right)=0\,,\end{aligned} \tag{A.5}$$

with similar components as derived above for $\frac{\partial W_1^{PB}}{\partial t_1}$. Again, from (A.5), we have $\left(t_2-(1-\gamma)\,D^{'}\right)\left(\frac{dx_{22}^{PB}}{dt_2}+\frac{dx_{21}^{PB}}{dt_2}\right)-(1-\gamma)\,D^{'}\left(\frac{dx_{11}^{PB}}{dt_2}+\frac{dx_{12}^{PB}}{dt_2}\right)$, and, hence, the $EDE_2=(1-\gamma)\,D^{'}+\frac{(1-\gamma)\,D^{'}\left(\frac{dx_{11}^{PB}}{dt_2}+\frac{dx_{12}^{PB}}{dt_2}\right)}{\Lambda_2^{PB}}$, where the term in brackets is the FEE_2.

Solving $\frac{\partial W_i^{PB}}{\partial t_i}$ for t_i gives the optimal PB-tax for country 1 and country 2 in (10) and (12), respectively.

A.4 BI-regime

The first-order conditions of profit-maximisation in (4) and (5) can be written as:

$\frac{\partial \pi_{11}}{\partial x_{11}} = p_1 + x_{11}p_1^{'} - c - t_1 = 0$ & $\frac{\partial \pi_{21}}{\partial x_{21}} = p_1 + x_{21}p_1^{'} - c - t_1 = 0\,,$

$\frac{\partial \pi_{12}}{\partial x_{12}} = p_2 + x_{12}p_2^{'} - c - t_1 = 0$ & $\frac{\partial \pi_{22}}{\partial x_{22}} = p_2 + x_{22}p_2^{'} - c - t_2 = 0\,.$

The effect of t_1 on market 1 follows from total differentiation of the first-order conditions. In market 1, we have:

$\frac{\partial^2 \pi_{11}}{\partial x_{11}^2} dx_{11} + \frac{\partial^2 \pi_{11}}{\partial x_{11}\partial x_{21}} dx_{21} + \frac{\partial^2 \pi_{11}}{\partial x_{11}\partial t_1} dt_1 = 0$, and

$\frac{\partial^2 \pi_{21}}{\partial x_{21}\partial x_{11}} dx_{11} + \frac{\partial^2 \pi_{21}}{\partial x_{21}^2} dx_{21} + \frac{\partial^2 \pi_{21}}{\partial x_{21}\partial t_1} dt_1 = 0\,.$

From Appendix A.1 and by substituting $\frac{\partial^2 \pi_{11}}{\partial x_{11}\partial t_1} = \frac{\partial^2 \pi_{21}}{\partial x_{21}\partial t_1} = -1$, we obtain:

$\begin{bmatrix} 2p_1^{'} + x_{11}p_1^{''} & p_1^{'} + x_{11}p_1^{''} \\ p_1^{'} + x_{21}p_1^{''} & 2p_1^{'} + x_{21}p_1^{''} \end{bmatrix} \begin{bmatrix} dx_{11} \\ dx_{21} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} dt_1,$

$\frac{dx_{11}^{BI}}{dt_1} = \begin{bmatrix} 1 & p_1^{'} + x_{11}p_1^{''} \\ 1 & 2p_1^{'} + x_{21}p_1^{''} \end{bmatrix} / J_1 = \frac{p_1^{'} + p_1^{''}(x_{21} - x_{11})}{J_1} = \frac{p_1^{'}}{J_1} < 0$, given $x_{11} = x_{21}$ and because both firms are identical and face the same effective tax.

$\frac{dx_{21}^{BI}}{dt_1} = \begin{bmatrix} 2p_1^{'} + x_{11}p_1^{''} & 1 \\ p_1^{'} + x_{21}p_1^{''} & 1 \end{bmatrix} / J_1 = \frac{p_1^{'} + p_1^{''}(x_{11} - x_{21})}{J_1} = \frac{p_1^{'}}{J_1} < 0$, and $\frac{dx_{11}^{BI}}{dt_1} + \frac{dx_{21}^{BI}}{dt_1} = \frac{2p_1^{'}}{J_1}$.

The effect of t_2 on market 1 is obviously zero.

The effect of t_1 and t_2 on market 2 is the same as under the PB-regime.

The welfare function of country 1 and 2 under this regime are given by:

$$W_1^{BI} = u_1(X_1) - p_1 X_1 + \Pi_1 + t_1(x_{11} + x_{12}) + (t_1 - t_2)x_{21} - D_1(e)\,, \tag{A.6}$$

$$W_2^{BI} = u_2(X_2) - p_2 X_2 + \Pi_2 + t_2(x_{22} + x_{21}) - D_2(e)\,. \tag{A.7}$$

The first-order condition from maximising A.6 is given by:

$$\begin{aligned} \frac{\partial W_1^{BI}}{\partial t_1} &= -p_1^{'}\left(\frac{dx_{11}^{BI}}{dt_1} + \frac{dx_{21}^{BI}}{dt_1}\right)(x_{11} + x_{21}) + p_1^{'}x_{11}\frac{dx_{21}^{BI}}{dt_1} + p_2^{'}x_{12}\frac{dx_{22}^{BI}}{dt_1} \\ &+ t_1\left(\frac{dx_{11}^{BI}}{dt_1} + \frac{dx_{12}^{BI}}{dt_1} + \frac{dx_{21}^{BI}}{dt_1}\right) + x_{21} - t_2\frac{dx_{21}^{BI}}{dt_1} - \gamma D^{'}\frac{\partial e}{\partial t_1} = 0\,. \end{aligned} \tag{A.8}$$

Rearranging terms, we have: $\frac{\partial W_1^{BI}}{\partial t_1} = -p_1^{'}\left(\frac{dx_{11}^{BI}}{dt_1} + \frac{dx_{21}^{BI}}{dt_1}\right)(x_{11} + x_{21}) + p_1^{'}x_{11}\frac{dx_{21}^{BI}}{dt_1} + p_2^{'}x_{12}\frac{dx_{22}^{BI}}{dt_1} + x_{21} - t_2\frac{dx_{21}^{BI}}{dt_1} + \left(t_1 - \gamma D^{'}\right)\left(\frac{dx_{11}^{BI}}{dt_1} + \frac{dx_{12}^{BI}}{dt_1} + \frac{dx_{21}^{BI}}{dt_1}\right) - \gamma D^{'}\left(\frac{dx_{22}^{BI}}{dt_1}\right) = 0$

and solving for t_1 gives the optimal carbon tax in (14) in the text. From the above equation, the EDE_1 under this regime can be written as $\gamma\, D^{'} + \frac{\gamma\, D^{'}\left(\frac{dx_{22}^{BI}}{dt_1}\right)}{\Lambda_1^{BI}}$.

The new component in this regime for country 1 is BCAI, where $\frac{\partial BCAI}{\partial t_1} = t_1 \frac{dx_{21}^{BI}}{dt_1} + x_{21} - t_2 \frac{dx_{21}^{BI}}{dt_1}$.

Under this regime, we have $\frac{\partial \Pi_1}{\partial t_1} = \underbrace{p_1^{'} x_{11} \frac{dx_{21}^{BI}}{dt_1}}_{+} + \underbrace{p_2^{'} x_{12} \frac{dx_{22}^{BI}}{dt_1}}_{-} - x_{11} - x_{12}$.

If the inverse demand curve is linear, the PSE is given by $-\frac{p_1^{'} p_2^{'} (x_{11} - x_{12})}{2p_1^{'} + 2p_2^{'}} > 0$, noting that $x_{11} > x_{12}$ and $t_1 > t_2$ because t_1 has a larger negative effect on x_{12} than on x_{11}, i.e., $\frac{dx_{11}^{BI}}{dt_1} = \frac{p_1^{'}}{J_1} < \frac{2p_2^{'}}{J_2} = \frac{dx_{12}^{BI}}{dt_1}$.

The first-order condition derived from maximisation of (A.7) is given by:

$$\frac{\partial W_2^{BI}}{\partial t_2} = -p_2^{'} \left(\frac{dx_{22}^{BI}}{dt_2} + \frac{dx_{12}^{BI}}{dt_2} \right) (x_{22} + x_{12}) + p_2^{'} x_{22} \frac{dx_{12}^{BI}}{dt_2} + t_2 \left(\frac{dx_{22}^{BI}}{dt_2} \right) + x_{21} - (1 - \gamma) D^{'} \frac{\partial e}{\partial t_2} = 0 \,. \tag{A.9}$$

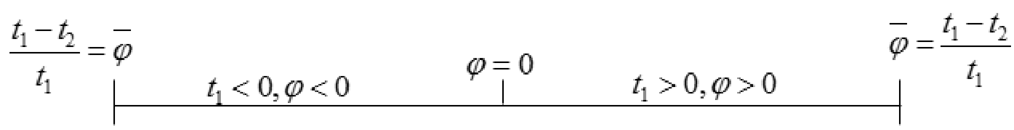
Rearranging terms, we have:

$\frac{\partial W_2^{BI}}{\partial t_2} = -p_2^{'} \left(\frac{dx_{22}^{BI}}{dt_2} + \frac{dx_{12}^{BI}}{dt_2} \right) (x_{22} + x_{12}) + p_2^{'} x_{22} \frac{dx_{12}^{BI}}{dt_2} + x_{21} + \left(t_2 - (1 - \gamma) D^{'} \right) \left(\frac{dx_{22}^{BI}}{dt_2} \right) - (1 - \gamma) D^{'} \left(\frac{dx_{12}^{BI}}{dt_2} \right) = 0$

and solving for t_2, gives the optimal tax in (16), where $\frac{dx_{11}^{BI}}{dt_2} = \frac{dx_{21}^{BI}}{dt_2} = 0$, and, hence $\frac{\partial \pi_{21}}{\partial t_2} = 0$. Therefore, $\frac{\partial \Pi_2}{\partial t_2} = \frac{\partial \pi_{22}}{\partial t_2} = p_2^{'} x_{22} \frac{dx_{12}^{BI}}{dt_2} - x_{22}$. In addition, $\frac{\partial TR_2}{\partial t_2} = t_2 \left(\frac{dx_{22}^{BI}}{dt_2} \right) + x_{22} + x_{21}$. Furthermore, from the above equation, it is clear that the EDE_2 becomes $(1 - \gamma)\ D^{'} + \frac{(1-\gamma)\, D^{'} \left(\frac{dx_{12}^{BI}}{dt_2} \right)}{\Lambda_2^{BI}}$.

A.5 BIE-regime

In models assuming imperfect competition, the equilibrium carbon tax can be positive or negative. Therefore, the feasible values of the rebate rate depends on the equilibrium policy in country 1 and 2. If $t_1 > 0$, $\varphi > 0$, and we have $0 < \varphi \leq \overline{\varphi} = \frac{t_1 - t_2}{t_1}$, where the maximum allowable rebate rate $\overline{\varphi}$ is $\overline{\varphi} \leq 1$ if $t_1 > t_2 \geq 0$, while $\overline{\varphi} > 1$ if $t_2 < 0$. However, if $0 > t_1 > t_2$, $\varphi < 0$. In this case, the feasible values for φ is $0 > \varphi \geq \overline{\varphi} = \frac{t_1 - t_2}{t_1}$. This can be illustrated as follows:



The first-order conditions of profit-maximisation in (4) and (5) can be written as:

$\frac{\partial \pi_{11}}{\partial x_{11}} = p_1 + x_{11}p_1^{'} - c - t_1 = 0$ & $\frac{\partial \pi_{21}}{\partial x_{21}} = p_1 + x_{21}p_1^{'} - c - t_1 = 0\,,$

$\frac{\partial \pi_{12}}{\partial x_{12}} = p_2 + x_{12}p_2^{'} - c - t_1(1-\varphi) = 0$ & $\frac{\partial \pi_{22}}{\partial x_{22}} = p_2 + x_{22}p_2^{'} - c - t_2 = 0\,.$

The effect of t_1 on market 1 and the effect of t_2 on market 1 and market 2 are the same as in the previous regime. The only change under this regime is the effect of t_1 on market 2:

$\frac{\partial^2 \pi_{12}}{\partial x_{12}^2} dx_{12} + \frac{\partial^2 \pi_{12}}{\partial x_{12}\partial x_{22}} dx_{22} + \frac{\partial^2 \pi_{12}}{\partial x_{12}\partial t_1} dt_1 = 0\,,$ and

$\frac{\partial^2 \pi_{22}}{\partial x_{22}\partial x_{12}} dx_{12} + \frac{\partial^2 \pi_{22}}{\partial x_{22}^2} dx_{22} + \frac{\partial^2 \pi_{22}}{\partial x_{22}\partial t_1} dt_1 = 0\,.$

From Appendix A.1 and by substituting $\frac{\partial^2 \pi_{12}}{\partial x_{12}\partial t_1} = -(1-\varphi)$ and $\frac{\partial^2 \pi_{22}}{\partial x_{22}\partial t_1} = 0$, we obtain:

$$\begin{bmatrix} 2p_2^{'} + x_{12}p_2^{''} & p_2^{'} + x_{12}p_2^{''} \\ p_2^{'} + x_{22}p_2^{''} & 2p_2^{'} + x_{22}p_2^{''} \end{bmatrix} \begin{bmatrix} dx_{12} \\ dx_{22} \end{bmatrix} = \begin{bmatrix} (1-\varphi) \\ 0 \end{bmatrix} dt_1\,,$$

$$\frac{dx_{12}^{BIE}}{dt_1} = \begin{bmatrix} (1-\varphi) & p_2^{'} + x_{12}p_2^{''} \\ 0 & 2p_2^{'} + x_{22}p_2^{''} \end{bmatrix} / J_2 = \frac{(1-\varphi)\left(2p_2^{'} + x_{22}p_2^{''}\right)}{J_2}\,, \text{ and}$$

$$\frac{dx_{22}^{BIE}}{dt_1} = \begin{bmatrix} 2p_2^{'} + x_{12}p_2^{''} & (1-\varphi) \\ p_2^{'} + x_{22}p_2^{''} & 0 \end{bmatrix} / J_2 = -\frac{(1-\varphi)\left(p_2^{'} + x_{22}p_2^{''}\right)}{J_2}\,.$$

In addition, the effects of the rebate rate on outputs, are given by:

$\frac{\partial^2 \pi_{12}}{\partial x_{12}^2} dx_{12} + \frac{\partial^2 \pi_{12}}{\partial x_{12}\partial x_{22}} dx_{22} + \frac{\partial^2 \pi_{12}}{\partial x_{12}\partial \varphi} d\varphi = 0\,,$ and

$\frac{\partial^2 \pi_{22}}{\partial x_{22}\partial x_{12}} dx_{12} + \frac{\partial^2 \pi_{22}}{\partial x_{22}^2} dx_{22} + \frac{\partial^2 \pi_{22}}{\partial x_{22}\partial \varphi} d\varphi = 0\,.$

From Appendix A.1 and by substituting $\frac{\partial^2 \pi_{12}}{\partial x_{12}\partial \varphi} = t_1$ and $\frac{\partial^2 \pi_{22}}{\partial x_{22}\partial \varphi} = 0$, we obtain:

$$\begin{bmatrix} 2p_2^{'} + x_{12}p_2^{''} & p_2^{'} + x_{12}p_2^{''} \\ p_2^{'} + x_{22}p_2^{''} & 2p_2^{'} + x_{22}p_2^{''} \end{bmatrix} \begin{bmatrix} dx_{12} \\ dx_{22} \end{bmatrix} = \begin{bmatrix} -t_1 \\ 0 \end{bmatrix} d\varphi\,,$$

$$\frac{dx_{12}^{BIE}}{d\varphi} = \begin{bmatrix} -t_1 & p_2^{'} + x_{12}p_2^{''} \\ 0 & 2p_2^{'} + x_{22}p_2^{''} \end{bmatrix} / J_2 = \frac{(-t_1)\left(2p_2^{'} + x_{22}p_2^{''}\right)}{J_2} \begin{cases} > 0 & if\, t_1 > 0 \\ < 0 & if\, t_1 < 0 \end{cases}\,, \text{ and}$$

$$\frac{dx_{22}^{BIE}}{d\varphi} = \begin{bmatrix} 2p_2^{'} + x_{12}p_2^{''} & -t_1 \\ p_2^{'} + x_{22}p_2^{''} & 0 \end{bmatrix} / J_2 = -\frac{(-t_1)\left(p_2^{'} + x_{22}p_2^{''}\right)}{J_2} \begin{cases} < 0 & if\, t_1 > 0 \\ > 0 & if\, t_1 < 0 \end{cases}\,,$$

where $\frac{dX_2}{d\varphi} = \frac{dx_{12}^{BIE}}{d\varphi} + \frac{dx_{22}^{BIE}}{d\varphi} = \frac{-t_1 p_2^{'}}{J_2} \begin{cases} > 0 & if\, t_1 > 0 \\ < 0 & if\, t_1 < 0 \end{cases}.$

Note that these effects can be explained in the case of a subsidy, i.e., $t_1 < 0$, as follows: if $t_1 < 0$, $\varphi < 0$ and hence higher φ means a lower subsidy and the exports of firm 1 decrease. The sign of the PSE_1 and the EDE_1 in (18) in the text will just be reversed if $t_1 < 0$.

The welfare function of country 1 and 2 are given by :

$$W_1^{BIE} = u_1(X_1) - p_1X_1 + \Pi_1 + t_1(x_{11} + x_{12}) + (t_1 - t_2)x_{21} - \varphi t_1 x_{12} - D_1(e)\ , \tag{A.10}$$

and

$$W_2^{BIE} = u_2(X_2) - p_2X_2 + \Pi_2 + t_2(x_{22} + x_{21}) - D_2(e)\ . \tag{A.11}$$

First, we derive the optimal rebate rate by maximising (A.10) with respect to φ:
$\frac{\partial W_1^{BIE}}{\partial \varphi} = \frac{\partial \pi_{12}}{\partial \varphi} + \frac{\partial TR_1}{\partial \varphi} - \frac{\partial BCAE}{\partial \varphi} - \frac{\partial D_1}{\partial \varphi} = 0$, or
$\frac{\partial W_1^{BIE}}{\partial \varphi} = p_2^{'}x_{12}\frac{dx_{22}^{BIE}}{d\varphi} + t_1x_{12} + t_1\frac{dx_{12}^{BIE}}{d\varphi} - \varphi t_1\frac{dx_{12}^{BIE}}{d\varphi} - t_1x_{12} - \gamma D'\left(\frac{dx_{12}^{BIE}}{d\varphi} + \frac{dx_{22}^{BIE}}{d\varphi}\right) = 0$
.

Solving the above first-order condition for φ leads to the optimal export rebate rate derived in (18), where $\frac{\partial \pi_{12}}{\partial \varphi} = (p_2 - c - t_1 + \varphi t_1)\left(\frac{dx_{12}}{d\varphi}\right) + x_{12}\left(p_2^{'}\left(\frac{dx_{12}}{d\varphi} + \frac{dx_{22}}{d\varphi}\right) + t_1\right)$.
From the first-order condition: $p_2 - c - t_1 + \varphi t_1 = -x_{12}p_2^{'}$. Therefore, $\frac{\partial \pi_{12}}{\partial \varphi} = x_{12}p_2^{'}\frac{dx_{22}}{d\varphi} + t_1x_{12}$,
$\frac{\partial TR_1}{\partial \varphi} = t_1\frac{dx_{12}}{d\varphi}$ and $\frac{\partial BCAE}{\partial \varphi} = \varphi t_1\frac{dx_{12}}{d\varphi} + t_1x_{12}$.

We now derive the optimal taxes of countries.

The first-order condition of (A.10) is given by:

$$\begin{aligned}\frac{\partial W_1^{BIE}}{\partial t_1} &= -p_1^{'}\left(\frac{dx_{11}^{BIE}}{dt_1} + \frac{dx_{21}^{BIE}}{dt_1}\right)(x_{11} + x_{21}) + p_1^{'}x_{11}\frac{dx_{21}^{BIE}}{dt_1} + p_2^{'}x_{12}\frac{dx_{22}^{BIE}}{dt_1} \\ &+ t_1\left(\frac{dx_{11}}{dt_1} + \frac{dx_{21}}{dt_1} + (1-\varphi)\frac{dx_{12}}{dt_1}\right) + x_{21} - t_2\frac{dx_{21}^{BIE}}{dt_1} - \gamma D^{'}\frac{\partial e}{\partial t_1} = 0\,.\end{aligned} \tag{A.12}$$

Rearranging terms lead to:
$\frac{\partial W_1^{BIE}}{\partial t_1} = -p_1^{'}\left(\frac{dx_{11}^{BIE}}{dt_1} + \frac{dx_{21}^{BIE}}{dt_1}\right)(x_{11} + x_{21}) + p_1^{'}x_{11}\frac{dx_{21}^{BIE}}{dt_1} + p_2^{'}x_{12}\frac{dx_{22}^{BIE}}{dt_1} + x_{21} - t_2\frac{dx_{21}^{BIE}}{dt_1} + \left(t_1 - \gamma D'\right)\left(\frac{dx_{11}^{BIE}}{dt_1} + \frac{dx_{21}^{BIE}}{dt_1} + (1-\varphi)\frac{dx_{12}^{BIE}}{dt_1}\right) - \gamma D'\left(\varphi\frac{dx_{12}^{BIE}}{dt_1} + \frac{dx_{22}^{BIE}}{dt_1}\right) = 0\,.$

Hence, from the above equation, the EDE_1 can be written as $\gamma D' + \frac{\gamma D'\left(\varphi\frac{dx_{12}^{BIE}}{dt_1} + \frac{dx_{22}^{BIE}}{dt_1}\right)}{\Lambda_1^{BIE}}$.

The new component in the welfare function of country 1 in this regime is the BCAE: $\frac{\partial BCAE}{\partial t_1} = \varphi t_1\frac{dx_{12}}{dt_1} + \varphi x_{12}\,.$

Solving the above condition for t_1 leads to the optimal tax in (20), where the effect of t_1 on the profits of firm 1 are:

$\frac{\partial \Pi_1}{\partial t_1} = p_1^{'} x_{11} \frac{dx_{21}^{BIE}}{dt_1} + p_2^{'} x_{12} \frac{dx_{22}^{BIE}}{dt_1} - x_{11} - x_{12} + \varphi x_{12}$.

The differentiation with respect to country 2 is the same as in the previous regime in (A.9), so the formula of the optimal tax in (22) is the same as in (16).

A.6 BF-regime

The first-order conditions of profit-maximisation in (4) and (5) can be written as:

$\frac{\partial \pi_{11}}{\partial x_{11}} = p_1 + x_{11} p_1^{'} - c - t_1 = 0$ & $\frac{\partial \pi_{21}}{\partial x_{21}} = p_1 + x_{21} p_1^{'} - c - t_1 = 0\,,$

$\frac{\partial \pi_{12}}{\partial x_{12}} = p_2 + x_{12} p_2^{'} - c = 0$ & $\frac{\partial \pi_{22}}{\partial x_{22}} = p_2 + x_{22} p_2^{'} - c - t_2 = 0$.

The effect of t_1 on market 1 and the effect of t_2 on market 1 and market 2 are the same as in the previous regimes. However, the effect of t_1 on market 2 is now:

$\frac{\partial^2 \pi_{12}}{\partial x_{12}^2} dx_{12} + \frac{\partial^2 \pi_{12}}{\partial x_{12} \partial x_{22}} dx_{22} + \frac{\partial^2 \pi_{12}}{\partial x_{12} \partial t_1} dt_1 = 0$ and

$\frac{\partial^2 \pi_{22}}{\partial x_{22} \partial x_{12}} dx_{12} + \frac{\partial^2 \pi_{22}}{\partial x_{22}^2} dx_{22} + \frac{\partial^2 \pi_{22}}{\partial x_{22} \partial t_1} dt_1 = 0$.

From Appendix A.1 and by substituting $\frac{\partial^2 \pi_{12}}{\partial x_{12} \partial t_1} = 0$ and $\frac{\partial^2 \pi_{22}}{\partial x_{22} \partial t_1} = 0$, we obtain:

$$\begin{bmatrix} 2p_2^{'} + x_{12} p_2^{''} & p_2^{'} + x_{12} p_2^{''} \\ p_2^{'} + x_{22} p_2^{''} & 2p_2^{'} + x_{22} p_2^{''} \end{bmatrix} \begin{bmatrix} dx_{12} \\ dx_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} dt_1\,,$$

$$\frac{dx_{12}^{BF}}{dt_1} = \begin{bmatrix} 0 & p_2^{'} + x_{12} p_2^{''} \\ 0 & 2p_2^{'} + x_{22} p_2^{''} \end{bmatrix} / J_2 = 0\,, \text{ and}$$

$$\frac{dx_{22}^{BF}}{dt_1} = \begin{bmatrix} 2p_2^{'} + x_{12} p_2^{''} & 0 \\ p_2^{'} + x_{22} p_2^{''} & 0 \end{bmatrix} / J_2 = 0\,.$$

The welfare function of country 1 and 2 are given by:

$$W_1^{BF} = u_1(X_1) - p_1 X_1 + \Pi_1 + t_1(x_{11} + x_{12}) + (t_1 - t_2) x_{21} - t_1 x_{12} - D_1(e)\ , \tag{A.13}$$

and

$$W_2^{BF} = u_2(X_2) - p_2 X_2 + \Pi_2 + t_2(x_{22} + x_{21}) - D_2(e)\ . \tag{A.14}$$

Maximising (A.13) with respect to t_1 leads to the following first-order condition:

$$\begin{aligned} \frac{\partial W_1^{BF}}{\partial t_1} = & -p_1^{'} \left(\frac{dx_{11}^{BF}}{dt_1} + \frac{dx_{21}^{BF}}{dt_1} \right) (x_{11} + x_{21}) + p_1^{'} x_{11} \frac{dx_{21}^{BF}}{dt_1} \\ & + x_{21} - t_2 \frac{dx_{21}}{dt_1} + \left(t_1 - \gamma D^{'} \right) \left(\frac{dx_{11}^{BF}}{dt_1} + \frac{dx_{21}^{BF}}{dt_1} \right) = 0\,. \end{aligned} \tag{A.15}$$

Solving this first-order condition for t_1 gives the optimal carbon tax of country 1 in (24), where $\frac{\partial \pi_{12}}{\partial t_1} = 0$, $\frac{\partial \Pi_1}{\partial t_1} = \frac{\partial \pi_{11}}{\partial t_1} = p_1^{'} x_{11} \frac{dx_{21}^{BF}}{dt_1} - x_{11}$, $\frac{\partial TR_1}{\partial t_1} = t_1 \left(\frac{dx_{11}^{BF}}{dt_1} \right) + x_{11} + x_{12}$ and $\frac{\partial BCAE}{\partial t_1} = x_{12}$, where $\frac{\partial TR_1}{\partial t_1} - \frac{\partial BCAE}{\partial t_1} = t_1 \left(\frac{dx_{11}}{dt_1} \right) + x_{11}$.

Simplifying the optimal tax structure gives:

$\hat{t}_1^{BF} = p_1^{'} x_{21} + \frac{p_1^{'} x_{11}\left(p_1^{'}/J_1\right)}{\left(2p_1^{'}/J_1\right)} - \frac{x_{21}}{\left(2p_1^{'}/J_1\right)} + \frac{t_2\left(p_1^{'}/J_1\right)}{\left(2p_1^{'}/J_1\right)} + \gamma D^{'}$, which leads to the simplified formula in (24), and under this regime the $EDE_1 = \gamma D^{'}$ as follows directly from (A.15).

The first-order condition of country 2 is the same as under the previous BCA-regimes, and, hence, also the optimal tax structure.

A.7 CB-regime

The first-order conditions of profit-maximisation in (4) and (5) can be written as:

$\frac{\partial \pi_{11}}{\partial x_{11}} = p_1 + x_{11} p_1^{'} - c - t_1 = 0$ & $\frac{\partial \pi_{21}}{\partial x_{21}} = p_1 + x_{21} p_1^{'} - c - t_1 = 0\,,$

$\frac{\partial \pi_{12}}{\partial x_{12}} = p_2 + x_{12} p_2^{'} - c - t_2 = 0$ & $\frac{\partial \pi_{22}}{\partial x_{22}} = p_2 + x_{22} p_2^{'} - c - t_2 = 0\,.$

The effect of t_1 on both markets are similar to the previous regime, hence the change is the effect of t_2 on market 2:

$\frac{\partial^2 \pi_{12}}{\partial x_{12}^2} dx_{12} + \frac{\partial^2 \pi_{12}}{\partial x_{12}\partial x_{22}} dx_{22} + \frac{\partial^2 \pi_{12}}{\partial x_{12}\partial t_2} dt_2 = 0\,,$ and

$\frac{\partial^2 \pi_{22}}{\partial x_{22}\partial x_{12}} dx_{12} + \frac{\partial^2 \pi_{22}}{\partial x_{22}^2} dx_{22} + \frac{\partial^2 \pi_{22}}{\partial x_{22}\partial t_2} dt_2 = 0\,.$

From Appendix A.1 and by substituting $\frac{\partial^2 \pi_{12}}{\partial x_{12}\partial t_2} = -1$ and $\frac{\partial^2 \pi_{22}}{\partial x_{22}\partial t_2} = -1$, we obtain:

$$\begin{bmatrix} 2p_2^{'} + x_{12} p_2^{''} & p_2^{'} + x_{12} p_2^{''} \\ p_2^{'} + x_{22} p_2^{''} & 2p_2^{'} + x_{22} p_2^{''} \end{bmatrix} \begin{bmatrix} dx_{12} \\ dx_{22} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} dt_2\ ,$$

$$\frac{dx_{12}^{CB}}{dt_2} = \begin{bmatrix} 1 & p_2^{'} + x_{12} p_2^{''} \\ 1 & 2p_2^{'} + x_{22} p_2^{''} \end{bmatrix} / J_2 = \frac{p_2^{'} + p_2^{''}(x_{22} - x_{12})}{J_2} = \frac{p_2^{'}}{J_2} < 0\,,$$

$$\frac{dx_{22}^{CB}}{dt_2} = \begin{bmatrix} 2p_2^{'} + x_{12} p_2^{''} & 1 \\ p_2^{'} + x_{22} p_2^{''} & 1 \end{bmatrix} / J_2 = \frac{p_2^{'} + p_2^{''}(x_{12} - x_{22})}{J_2} = \frac{p_2^{'}}{J_2} < 0\,,$$

and $\frac{dx_{12}^{CB}}{dt_2} + \frac{dx_{22}^{CB}}{dt_2} = \frac{2p_2^{'}}{J_2} < 0.$

The welfare function of each country under this regime is given by:

$$W_1^{CB} = u_1\left(X_1\right) - p_1 X_1 + \Pi_1 + t_1\left(x_{11} + x_{21}\right) - D_1\left(e\right)\ , \tag{A.16}$$

$$W_2^{CB} = u_2\left(X_2\right) - p_2 X_2 + \Pi_2 + t_2\left(x_{22} + x_{12}\right) - D_2\left(e\right)\ . \tag{A.17}$$

Maximising (A.16) with respect to t_1 gives:

$$\frac{\partial W_1^{CB}}{\partial t_1} = -p_1^{'}\left(\frac{dx_{11}}{dt_1}+\frac{dx_{21}}{dt_1}\right)(x_{11}+x_{21}) + p_1^{'}x_{11}\frac{dx_{21}}{dt_1}$$
$$+\left(t_1-\gamma D^{'}\right)\left(\frac{dx_{11}}{dt_1}+\frac{dx_{21}}{dt_1}\right)+x_{21}=0\,. \tag{A.18}$$

Similarly, maximising (A.17) with respect to t_2, delivers:

$$\frac{\partial W_2^{CB}}{\partial t_2} = -p_2^{'}\left(\frac{dx_{12}}{dt_2}+\frac{dx_{22}}{dt_2}\right)(x_{12}+x_{22}) + p_2^{'}x_{22}\frac{dx_{12}}{dt_2}$$
$$+\left(t_2-(1-\gamma)\,D^{'}\right)\left(\frac{dx_{12}}{dt_2}+\frac{dx_{22}}{dt_2}\right)+x_{12}=0\,. \tag{A.19}$$

Solving the above first-order condition for t_i, gives the optimal CB-tax of country 1 and country 2 in (28) and (30), respectively. The simplified formula can be obtained in the same manner as under the BF-regime in the previous appendix, even though the term $\frac{t_2\left(p_1^{'}/J_1\right)}{\left(2p_1^{'}/J_1\right)}$ disappears because country 2 does not tax its exports under the CB-regime.

B Specific Functions

From the linear demand function in (32) in Section 5, we have $p_i^{'}=-1$ and $p_i^{''}=0$. Hence, $J_i=3$. For the damage function in (33), we have $D'=de$.

The outcome of the second stage is given by:

$$x_{1i}^{*}=\frac{A-2t_{1i}+t_{2i}}{3}\ \&\ x_{2i}^{*}=\frac{A-2t_{2i}+t_{1i}}{3}\ \forall\ i=1,2\,, \tag{A.20}$$

with $A=a-c>0$, which we interpret as a market size, or, as a proxy of the net benefits of production and consumption.

In the first stage, taking the above equilibrium output levels as a given, the individual welfare functions are given by:

$$W_1=\underbrace{\frac{(x_{11}+x_{21})^2}{2}}_{CS_1}+\underbrace{(x_{11})^2+(x_{12})^2}_{PS_1}+\underbrace{t_1(x_{11}+x_{12})}_{TR_1}-\underbrace{\gamma d\frac{e^2}{2}}_{D_1}$$
$$+\underbrace{(t_1-t_2)x_{21}}_{BCAI_1}-\underbrace{\varphi t_1 x_{12}}_{BCAE_1}\,, \tag{A.21}$$

$$W_2 = \underbrace{\frac{(x_{22}+x_{12})^2}{2}}_{CS_2} + \underbrace{(x_{22})^2 + (x_{21})^2}_{PS_2} + \underbrace{t_2(x_{22}+x_{21})}_{TR_2} - \underbrace{(1-\gamma)\, d\frac{e^2}{2}}_{D_2}\,. \tag{A.22}$$

Note that under the CB-regime, $TR_i = t_i\,(x_{1i} + x_{2i})$.

Countries choose their tax levels cooperatively (in the social optimum) or non-cooperatively under the other regimes as explained in the text. The equilibrium tax levels are then inserted in the equilibrium output levels in (A.20). The feasible range for the parameters values is determined by two constraints: a non-negativity (NN) constraint, which ensures positive production levels by both firms in the two markets. This constraint implies that the damage parameter d is not too large, i.e., $d < \overline{d}\,(\gamma)$. In addition, because we assume that BCA-measures are imposed by country 1, a BCA-constraint is needed to ensure that $t_1 > t_2$ and $t_1(1-\varphi) \geq t_2$. The BCA-constraint implies that the damage parameter d is not too small, i.e., $d > \underline{d}\,(\gamma)$.

B.1. Social Optimum

Inserting the effective tax t_S into (A.20) gives equilibrium outputs $x_{ki} = \frac{A-t_S}{3}$. Inserting equilibrium outputs into the aggregate welfare function $W^S = W_1 + W_2$ (see (A.1)), implies the following first- and second order condition:

$\frac{\partial W^S}{\partial t_S} = \frac{16d(A-t_S)-8t_S-4A}{9} = 0$, and $\frac{\partial^2 W}{\partial t_S^2} = -\frac{8}{9}\,(1+2d) < 0$.

Solving the first-order condition for t_S leads to:

$$t_S^* = \frac{A\,(4d-1)}{2+4d}\,. \tag{A.23}$$

Consequently, $x_{ki}^{*S} = A/\,(2+4d)$, where no NN-constraint is needed.

B.2. PB-regime

From Table 1, inserting effective taxes into (A.20), we obtain equilibrium outputs: $x_{11}^{PB} = x_{12}^{PB} = \frac{A-2t_1+t_2}{3}$ and $x_{22}^{PB} = x_{21}^{PB} = \frac{A-2t_2+t_1}{3}$. Inserting outputs into (A.21) and (A.22), the first-order conditions are given by:

$\frac{\partial W_1^{PB}}{\partial t_1} = \frac{1}{9}\left(8\gamma d\left(A - \frac{t_1}{2} - \frac{t_2}{2}\right) - 4A - 7t_1 - t_2\right) = 0$, and
$\frac{\partial W_2^{PB}}{\partial t_2} = \frac{1}{9}\left(8\,(1-\gamma)\,d\left(A - \frac{t_1}{2} - \frac{t_2}{2}\right) - 4A - 7t_2 - t_1\right) = 0.$

The second-order conditions are satisfied: $\frac{\partial^2 W_1}{\partial {t_1}^2} = -\frac{7}{9} - \frac{4\gamma d}{9} < 0$ and $\frac{\partial^2 W_2}{\partial {t_2}^2} = -\frac{7}{9} - \frac{4(1-\gamma)d}{9} < 0$. Moreover, we have $\frac{\partial^2 W_1}{\partial t_1 \partial t_2} = -\frac{1}{9} - \frac{4\gamma d}{9} < 0$ and $\frac{\partial^2 W_2}{\partial t_2 \partial t_1} = -\frac{1}{9} - \frac{4(1-\gamma)d}{9} < 0$, and, hence, $\frac{\partial^2 W_1}{\partial t_1{}^2}\frac{\partial^2 W_2}{\partial t_2{}^2} - \frac{\partial^2 W_1}{\partial t_1 \partial t_2}\frac{\partial^2 W_2}{\partial t_2 \partial t_1} = \frac{16+8d}{27} > 0$, which ensures a unique and stable equilibrium.

Solving the above two first-order conditions simultaneously, the equilibrium PB-taxes are given by:

$$t_1^{PB*} = \frac{A\left(4\gamma d - d - 1\right)}{d+2}, \tag{A.24}$$

$$t_2^{PB*} = \frac{A\left(3d - 4\gamma d - 1\right)}{d+2}. \tag{A.25}$$

Inserting these taxes into equilibrium outputs shows that the NN-constraint under this regime is given by $d < \frac{1}{2(2\gamma-1)}$ if $\gamma > 0.5$ and no constraint is needed if $\gamma = 0.5$.

B.3. BI-regime

Equilibrium outputs are given by: $x_{11}^{BI} = x_{21}^{BI} = \frac{A-t_1}{3}$, $x_{12}^{BI} = \frac{A-2t_1+t_2}{3}$ and $x_{22}^{BI} = \frac{A-2t_2+t_1}{3}$ (see Table 1). By inserting outputs into (A.21) and (A.22), the first-order condition can be derived:

$\frac{\partial W_1^{BI}}{\partial t_1} = \gamma d\left(\frac{4A-t_2}{3} - t_1\right) - A - \frac{10}{9}t_1 + \frac{2}{9}t_2 = 0$, and

$\frac{\partial W_2^{BI}}{\partial t_2} = \frac{1}{3}\left((1-\gamma)\, d\left(\frac{4A-t_2}{3} - t_1\right) - t_1 - t_2\right) = 0$,

where the second-order conditions hold: $\frac{\partial^2 W_1}{\partial t_1{}^2} = -\frac{10}{9} - \gamma d < 0$ and $\frac{\partial^2 W_2}{\partial t_2{}^2} = -\frac{1}{3} - \frac{(1-\gamma)d}{9} < 0$. Because $\frac{\partial^2 W_1}{\partial t_1 \partial t_2} = \frac{2}{9} - \frac{\gamma d}{3}$ and $\frac{\partial^2 W_2}{\partial t_2 \partial t_1} = -\frac{1}{3} - \frac{(1-\gamma)d}{3} < 0$, we have $\frac{\partial^2 W_1}{\partial t_1{}^2}\frac{\partial^2 W_2}{\partial t_2{}^2} - \frac{\partial^2 W_1}{\partial t_1 \partial t_2}\frac{\partial^2 W_2}{\partial t_2 \partial t_1} = \frac{4}{9} + \frac{(2\gamma+16d)}{81} > 0$, which ensures a unique and stable equilibrium.

Solving the first-order conditions, equilibrium carbon taxes are given by:

$$t_1^{BI*} = \frac{A\left(29\gamma d + 7d - 3\right)}{36 + 2d\left(\gamma + 8\right)}, \tag{A.26}$$

$$t_2^{BI*} = \frac{A\left(43d - 79\gamma d + 3\right)}{36 + 2d\left(\gamma + 8\right)}. \tag{A.27}$$

It can be shown that the most restrictive NN-constraint requires $d < \frac{1}{3\gamma-1}$. Since the difference between the two national tax levels is ambiguous, we need to impose a constraint on the parameters such that $t_1^{BI*} > t_2^{BI*}$, which requires the BCA-constraint $d > \frac{1}{6(3\gamma-1)}$.

B.4. BIE-regime

Equilibrium output levels are: $x_{11}^{BIE} = x_{21}^{BIE} = \frac{A-t_1}{3}$, $x_{12}^{BIE} = \frac{A-2t_1(1-\varphi)+t_2}{3}$ and $x_{22}^{BIE} = \frac{A-2t_2+t_1(1-\varphi)}{3}$.

Consequently, the first-order conditions are given by:

$\frac{\partial W_1^{BIE}}{\partial t_1} = \frac{1}{3}4\gamma d\left(A-\frac{3}{4}t_1-\frac{1}{4}t_2\right) - \frac{1}{9}\left(A+10t_1-2t_2\right) + \frac{1}{9}\varphi^2\left(-\gamma dt_1-4t_1\right) + \frac{1}{9}\varphi\left(A+8t_1+t_2-4\gamma d(A-\frac{3}{2}t_1-\frac{1}{4}t_2\right)=0\,,$

$\frac{\partial W_1^{BIE}}{\partial \varphi} = \frac{1}{3}t_1\left(\frac{1}{3}\left(A-2t_1\left(1-\varphi\right)+t_2\right)\right)+\frac{2}{3}t_1^2\left(1-\varphi\right)-\frac{1}{3}t_1\gamma d\left(\frac{4}{3}A-t_1+\frac{1}{3}t_1\varphi-\frac{1}{3}t_2\right)= 0\,,$ and

$\frac{\partial W_2^{BIE}}{\partial t_2} = \frac{1}{3}\left(\left(1-\gamma\right)d\left(\frac{4A-t_2}{3}-t_1+\frac{1}{3}t_1\varphi\right)-t_1-t_2\right)=0.$

Solving the first order conditions simultaneously, we have:

$$t_1^{BIE*} = \frac{17Ad\left(\gamma+1\right)}{36+d\left(5\gamma+14\right)} > 0\,, \tag{A.28}$$

$$t_2^{BIE*} = \frac{68Ad\left(1/2-\gamma\right)}{36+d\left(5\gamma+14\right)} \leq 0\,, \tag{A.29}$$

$$\varphi^* = \frac{d\left(29-37\gamma\right)+9}{17d\left(\gamma+1\right)} > 0\,, \tag{A.30}$$

noting that $\frac{\partial^2 W_1}{\partial {t_1}^2} = \frac{1}{9}\left(\varphi\left(6\gamma d+8\right)+\varphi^2\left(-\gamma d-4\right)-10\right)-\gamma d < 0 \forall\, \varphi$ and $\gamma, d>0$, $\frac{\partial^2 W_1}{\partial t_1 \partial t_2} = \frac{1}{9}\left(2+\varphi+\gamma d\left(\varphi-3\right)\right)$, $\frac{\partial^2 W_2}{\partial {t_2}^2} = -\frac{1}{3}-\frac{(1-\gamma)d}{9} < 0$, $\frac{\partial^2 W_2}{\partial t_2 \partial t_1} = -\frac{1}{3}+\frac{(1-\gamma)d(\varphi-3)}{9}$, and $\frac{\partial^2 W_1}{\partial \varphi^2} = -\frac{4}{9}t_1^2-\frac{1}{9}\gamma d t_1^2 < 0 \;\forall\, t_1 \neq 0$. Hence, $\frac{\partial^2 W_1}{\partial {t_1}^2}\frac{\partial^2 W_2}{\partial {t_2}^2}-\frac{\partial^2 W_1}{\partial t_1 \partial t_2}\frac{\partial^2 W_2}{\partial t_2 \partial t_1} > 0 \;\forall\; 0 \leq \gamma \leq 1$ and $d>0$.

Inserting equilibrium taxes into outputs, it turns out that the most restrictive NN-constraint requires $d < \frac{6}{19\gamma-8}$. We also need to impose a BCA-constraint such that $t_1^{BIE*}(1-\varphi^*) \geq t_2^{BIE*}$ or, equivalently, $\varphi^* \leq \frac{t_1^{BIE*}-t_2^{BIE*}}{t_1^{BIE*}}$, which leads to $d \geq \frac{9}{2(61\gamma-23)}$. $\varphi^*>0$ follows directly for all $d>0$ and $\gamma \leq \frac{29}{37} \simeq 0.78$, but is also true for $\gamma > 0.78$ if $d < \frac{9}{37\gamma-29}$, which holds due to the NN-constraint. Hence, $\varphi^* > 0$ is always true.

B.5. BF-regime

Equilibrium outputs are: $x_{11}^{BF} = x_{21}^{BF} = \frac{A-t_1}{3}$, $x_{12}^{BF} = \frac{A+t_2}{3}$ and $x_{22}^{BF} = \frac{A-2t_2}{3}$. Inserting these outputs into (A.21) and (A.22), the first-order conditions are:

$\frac{\partial W_1^{BF}}{\partial t_1} = \frac{1}{3}\left(-2t_1+t_2+2\gamma d\left(\frac{4A-2t_1-t_2}{3}\right)\right)=0\,,$ and
$\frac{\partial W_2^{BF}}{\partial t_2} = \frac{1}{3}\left(-t_1-t_2+\left(1-\gamma\right)d\left(\frac{4A-2t_1-t_2}{3}\right)\right)=0.$

The second derivatives are: $\frac{\partial^2 W_1}{\partial {t_1}^2} = -\frac{2}{3}-\frac{4\gamma d}{9} < 0$, $\frac{\partial^2 W_1}{\partial t_1 \partial t_2} = \frac{1}{3}-\frac{2\gamma d}{9}$, $\frac{\partial^2 W_2}{\partial {t_2}^2} = -\frac{1}{3}-\frac{(1-\gamma)d}{9} < 0$ and $\frac{\partial^2 W_2}{\partial t_2 \partial t_1} = -\frac{1}{3}-\frac{2(1-\gamma)d}{9} < 0$. Hence, $\frac{\partial^2 W_1}{\partial {t_1}^2}\frac{\partial^2 W_2}{\partial {t_2}^2}-\frac{\partial^2 W_1}{\partial t_1 \partial t_2}\frac{\partial^2 W_2}{\partial t_2 \partial t_1} = \frac{1}{3}+\frac{d(4-2\gamma)}{27} > 0 \,\forall\, \gamma \leq 2$.

Solving the first-order conditions, equilibrium taxes are given by:

$$t_1^{BF*} = \frac{-4Ad\,(\gamma+1)}{d\,(2\gamma-4)-9} > 0\,, \tag{A.31}$$

$$t_2^{BF*} = \frac{8Ad\,(2\gamma-1)}{d\,(2\gamma-4)-9} \leq 0\,. \tag{A.32}$$

Inserting equilibrium taxes into outputs, it turns out that the most restrictive NN-constraint requires $d < \frac{3}{2\gamma}$. There is no need for a BCA-constraint since $t_1^{BF*} > t_2^{BF*}$ and $t_1^{BF*}(1-\varphi) = 0 \geq t_2^{BF*}$ always hold.

B.6. CB-regime

Equilibrium outputs are: $x_{11}^{CB} = x_{21}^{CB} = \frac{A-t_1}{3}$, $x_{12}^{CB} = x_{22}^{CB} = \frac{A-t_2}{3}$. The first-order conditions are given by:

$\frac{\partial W_1^{CB}}{\partial t_1} = \frac{1}{3}\left(-2t_1 + 2\gamma d\left(\frac{4A-2t_1-2t_2}{3}\right)\right) = 0$, and

$\frac{\partial W_2^{CB}}{\partial t_2} = \frac{1}{3}\left(-2t_2 + 2\,(1-\gamma)\,d\left(\frac{4A-2t_1-2t_2}{3}\right)\right) = 0$. For the second derivatives we find: $\frac{\partial^2 W_1}{\partial {t_1}^2} = -\frac{2}{3} - \frac{4\gamma d}{9} < 0$, $\frac{\partial^2 W_1}{\partial t_1 \partial t_2} = -\frac{4\gamma d}{9} < 0$, $\frac{\partial^2 W_2}{\partial {t_2}^2} = -\frac{2}{3} - \frac{4(1-\gamma)d}{9} < 0$, and $\frac{\partial^2 W_2}{\partial t_2 \partial t_1} = -\frac{4(1-\gamma)d}{9} < 0$. Hence, $\frac{\partial^2 W_1}{\partial {t_1}^2}\frac{\partial^2 W_2}{\partial {t_2}^2} - \frac{\partial^2 W_1}{\partial t_1 \partial t_2}\frac{\partial^2 W_2}{\partial t_2 \partial t_1} = \frac{4}{9} + \frac{8d}{27} > 0$.

Solving the first-order conditions yields the CB-taxes:

$$t_1^{CB*} = \frac{4A\gamma d}{2d+3} = D_1^{'}(e)\,, \tag{A.33}$$

$$t_2^{CB*} = \frac{4A(1-\gamma)d}{2d+3} = D_2^{'}(e)\,. \tag{A.34}$$

Inserting equilibrium taxes into outputs, the NN-constraint requires $d < \frac{3}{4\gamma-2}$ if $\gamma > 0.5$ and if $\gamma = 0.5$ no constraint is required.

B.7. Ranges of Parameter Values

We summarise the conditions that satisfy the NN-constraint and the BCA-constraint in the following table and restrict our feasible range of parameter values to the most strict condition if we conduct a comparison across regimes.

Table A.1: Feasible Range of the Parameters Values

Regime/Constraint	NN-constraint	BCA-constraint
Social Optimum	/	/
PB	$d < \frac{1}{2(2\gamma-1)}$	/
BI	$d < \frac{1}{3\gamma-1}$	$d > \frac{1}{6(3\gamma-1)}$
BIE	$d < \frac{6}{19\gamma-8}$	$d \geq \frac{9}{2(61\gamma-23)}$
BF	$d < \frac{3}{2\gamma}$	/
CB	$d < \frac{3}{4\gamma-2}$	/
Feasible Range	$d < \overline{d}(\gamma) = \frac{1}{3\gamma-1}$	$d \geq \underline{d}(\gamma) = \frac{9}{2(61\gamma-23)}$

B.8. Proof of Proposition 5

Using the equilibrium tax levels from (A.23) to (A.34) and the feasible ranges for the parameter values from Table A.1, we conduct the following comparisons.

1. For country 1, (a) we have $t_1^{BI*} > t_1^{PB*}$ for all $\gamma \leq 0.607$ and if $d < \frac{\sqrt{8016\gamma^2-6624\gamma+1209}+63-84\gamma}{2(8\gamma^2+33\gamma-23)}$ for all $\gamma > 0.607$, where the NN-constraint is sufficient for this condition to hold. $t_1^{BIE*} > t_1^{BI*}$ for all $\gamma \leq \frac{29}{37} \approx 0.783$ and if $d < \frac{9}{37\gamma-29}$ for all $\gamma > 0.783$, where also this condition holds as long as the NN-constraint holds. In addition, $t_1^{BF*} > t_1^{BI*}$ for all $d > 0$. Finally, $t_1^{BF*} \geq t_1^{BIE*}$ if $d \geq \frac{3}{2(9\gamma-2)}$, which implies that $\varphi^* \leq 1$. However, if $d < \frac{3}{2(9\gamma-2)}$, $\varphi^* > 1$, and we have $t_1^{BIE*} > t_1^{BF*}$.

   (b) $t_1^{CB*} > t_1^{BF*}$ if $d < \frac{3(2\gamma-1)}{2(\gamma^2-\gamma+1)}$, where the NN-constraint is sufficient to guarantee this condition for all $\gamma \geq \mathring{\gamma} = 0.73$ which leads to the ranking in Proposition 5. Furthermore, $t_1^{CB*} > t_1^{BIE*}$ for all $\gamma \geq 0.865$ and if $d < -\frac{3(31\gamma-17)}{2(10\gamma^2+11\gamma-17)}$ for all $\gamma < 0.865$. However, also the NN-constraint is sufficient to guarantee this condition for all $0.66 \leq \gamma < 0.865$. In addition, this ranking also holds if $\gamma \in [0.61, 0.73)$ if $d < \frac{3(2\gamma-1)}{2(\gamma^2-\gamma+1)}$. However, if $\gamma < 0.61$, the tax under BCA-regimes may be larger than the CB-tax.

2. For country 2, (a) we have $t_2^{BI*} > t_2^{BIE*}$ if $d > \frac{9}{37\gamma-29} < 0$ for all $\gamma < \frac{29}{37} \approx 0.78$, and if $\gamma = 0.78$, and if $d < \frac{9}{37\gamma-29}$ for all $\gamma > 0.78$ which must hold due to the NN-constraint. In addition, $t_2^{BI*} > t_2^{BF*}$ is always true for all $d > 0$. Hence, $t_2^{BI*} > t_2^{BIE*}, t_2^{BF*}$. We also have $t_2^{BF*} \geq t_2^{BIE*}$ if $d \leq \frac{3}{2(9\gamma-2)}$, which implies that $\varphi^* \geq 1$. If $d > \frac{3}{2(9\gamma-2)}$, i.e. $\varphi^* < 1$, we have $t_2^{BIE*} > t_2^{BF*}$. Note that $t_2^{BIE*} = t_2^{BF*} = 0$ if $\gamma = 0.5$, irrespective of φ^*.

   In addition, $t_2^{BI*} > t_2^{CB*}$ if $d < -\frac{3\left(\sqrt{929\gamma^2+594\gamma-79}-31\gamma-3\right)}{4(4\gamma^2-51\gamma+11)}$, which violates the BCA-constraint for all values of γ. Hence, we have $t_2^{BI*} < t_2^{CB*}$, which leads to the ranking in (a).

(b) Comparing the equilibrium tax level under the PB-regime with the BCA-regimes leads to the following:

i. $t_2^{BI*} > t_2^{PB*}$ if $d < \breve{d} = -\frac{\sqrt{3600\gamma - 1200\gamma^2 + 849} - 12\gamma - 3}{2(8\gamma^2 - 21\gamma - 5)}$. The NN-constraint is sufficient for this condition to hold for all $\gamma > 0.57$, while if $d \geq \breve{d}$ for all $\gamma \leq 0.57$, we have $t_2^{BI*} \leq t_2^{PB*}$.

ii. $t_2^{BIE*} > t_2^{PB*}$ if $d < \dddot{d} = -\frac{\sqrt{3212\gamma - 2711\gamma^2 + 1828} - 13\gamma - 26}{2(20\gamma^2 - 27\gamma - 8)}$. This condition must hold for all $\gamma \geq 0.66$ as long the NN-constraint, i.e. $d < \overline{d}(\gamma)$, holds, while we could have $t_2^{BIE*} \leq t_2^{PB*}$ if $d \geq \dddot{d}$ for all $\gamma < 0.66$.

ii. $t_2^{BF*} > t_2^{PB*}$ if $d < \ddot{d} = \frac{2\gamma - 7 + \sqrt{292\gamma^2 - 244\gamma + 193}}{4(4\gamma^2 - 3\gamma + 2)}$, where the NN-constraint, $d < \overline{d}(\gamma)$, is sufficient for this condition to hold for all $\gamma > \ddot{\gamma} = 0.673$, while if $\gamma \leq \ddot{\gamma}$, we could have $t_2^{BF*} \leq t_2^{PB*}$ if $d \geq \ddot{d}$.

We have $\breve{d} > \dddot{d} > \ddot{d}$.

Therefore, we choose the most strict condition, i.e. $d < \ddot{d}$ as a sufficient condition for the ranking provided in Proposition 5.

3. $t_S^* > t_1^{PB*}$ if $\gamma = 0.5$ and if $d < -\frac{8\gamma - 13}{8(2\gamma - 1)}$ for all $\gamma > 0.5$. The NN-constraint (either the most restrictive or the one under the PB-regime only) is sufficient for this condition to hold. Similarly, $t_S^* > t_2^{PB*}$ if $\gamma = 0.5$ and if $d > -\frac{5 + 8\gamma}{16\gamma - 8}$ for all $\gamma > 0.5$ which must hold as the right hand side term is negative and d is a positive parameter.

   $t_1^{CB*} > t_1^{PB*}$ if $d < \frac{3}{4\gamma - 2}$ for all $\gamma > 0.5$. This is exactly the NN-constraint under the CB-regime (see table A.1), and, hence, must hold. For $\gamma = 0.5$, $t_1^{PB*} < t_1^{CB*}$ is easily checked. In addition, $t_2^{CB*} > t_2^{PB*}$ if $d > -\frac{3}{4\gamma - 2}$ for all $\gamma > 0.5$. which must hold as d is a positive parameter. Again, for $\gamma = 0.5$, $t_2^{CB*} > t_2^{PB*}$ is easily checked.

4. In order to obtain a sufficient condition for the ranking in the fourth result in Proposition 5, we need to compare the lowest non-cooperative tax level with the socially optimal tax. Since taxes of country 1 are always larger than that of country 2 under all regimes, we will choose the lowest tax level country 2 could set to compare with the socially optimal tax. It turns out, that the most strict condition follows from: $t_2^{BIE*} > t_S^*$ if $d < \tilde{d}(\gamma) = \frac{\sqrt{17161\gamma^2 + 58292\gamma - 7676} - 131\gamma - 62}{8(73\gamma - 20)}$. However, this condition violates the BCA-constraint for all $\gamma < \tilde{\gamma} \approx 0.77$, while for $\gamma \geq \tilde{\gamma}$, we need $d < \tilde{d}(\gamma)$, which leads to the ranking in Proposition 5. Note that for all $\gamma \geq \tilde{\gamma}$, $t_1^{CB*} > t_1^{BIE*}, t_1^{BF*}$ as shown above.

Part IV

Essay 3: Enforcing Climate Agreements: The Role of Escalating Border Carbon Adjustments

Appendix 6B: Statement of Authorship

This declaration concerns the article entitled:

Enforcing Climate Agreements: The Role of Escalating Border Carbon Adjustments

Publication status (tick one)

Draft manuscript	X	**Submitted**		**In review**		**Accepted**		**Published**	

Publication details (reference)	

Copyright status (tick the appropriate statement)

I hold the copyright for this material	X	Copyright is retained by the publisher, but I have been given permission to replicate the material here	

Candidate's contribution to the paper (provide details, and also indicate as a percentage)	The candidate contributed to / considerably contributed to / predominantly executed the… Formulation of ideas: - Predominantly contributed to the formulation of ideas. (65 %) Design of methodology: - Predominantly contributed to the design of methodology. (70 %) ~~Experimental work:~~ Presentation of data in journal format: - Predominantly contributed to the presentation of data in journal format. (65 %)
Statement from Candidate	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature.

Signed	Noha Nagi Elboghdadly	**Date**	4/10/2019

Enforcing Climate Agreements: The Role of Escalating Border Carbon Adjustments

Noha Elboghdadly* and Michael Finus†

Abstract

Border carbon adjustments (BCAs) have been suggested to reduce carbon leakage in the presence of unilateral climate policies and to enforce cooperative climate agreements. In an intra-industry trade model with two asymmetric countries, this paper studies whether and under which conditions an escalating sequence of BCA-measures could be successful in enforcing a fully cooperative agreement. We start from the assumption that moving from non-cooperative and non-uniform production-based carbon taxes to a uniform socially optimal tax is not attractive to the environmentally less concerned country. We then test whether the threat to impose BCA-measures, in the form of import tariffs only or complemented by export rebates, will enforce cooperation. We show that import tariffs are the least distortionary policy instrument but the weakest punishment, and import tariffs with a full export rebate is the most distortionary instrument if implemented but the harshest punishment to enforce cooperation. In an escalating penalty game, we determine the subgame-perfect equilibrium path in which threats must be deterrent but also credible. We show that BCA-measures help to enforce cooperation and are welfare improving if they need to be implemented. However, whenever full cooperation is be expected to generate the highest global welfare gains, BCAs fail to establish cooperation, a version of the paradox of cooperation, as proposed by Barrett (1994), though in a model without trade and BCAs.

Keywords: Border Carbon Adjustments, Escalating Penalties, Enforcement of Cooperation, Carbon Leakage

JEL-Classification: C7, F12, F18, Q58, H23

*Department of Economics, University of Bath, 3 East, Bath, BA2 7AY, UK. Email: n.m.w.elboghdadly@bath.ac.uk

†Department of Economics, University of Graz, Universitätsstraße 15, 8010 Graz, Austria and University of Bath, 3 East, Bath, BA2 7AY, UK. Email: michael.finus@uni-graz.at

1 Introduction

An effective solution to climate change requires cooperation among all countries in reducing global greenhouse gas emissions. However, strong free-rider incentives constitute a major stumbling block to reach a global agreement. Furthermore, the effectiveness of unilateral actions by some environmentally friendly countries is weakened by carbon leakage. One of the channels of carbon leakage is through the relocation of the production of firms, in particular in emission-intensive trade-exposed (EITE) industries, to countries with less strict climate policies. As a result, firms operating in countries with a stricter climate policy lose market shares in the domestic and international markets.

In order to both address carbon leakage and incentivise higher carbon taxes globally, border carbon adjustments (BCAs) have been proposed (Elliott et al., 2010; Helm et al., 2012; Ismer and Neuhoff, 2007; Stiglitz, 2006). BCAs are trade measures complementing climate policies. These measures are imposed against less regulated countries and may take the form of (i) BCAs on imports whereby a carbon tariff is levied on imports or (ii) BCAs on exports whereby a rebate is given on exports or (iii) both BCAs on imports and exports which is sometimes referred to as full BCAs.[1]

Trade measures to support environmental policies can be defended in theory, based on economic efficiency, as a second-best solution that corrects distortions resulting from the failure of internalising damages of transboundary pollution (Copeland, 1996; Helm et al., 2012; Hoel, 1996; Markusen, 1975; Tsakiris et al., 2014; Yonezawa et al., 2012). In the absence of global action, e.g., Markusen (1975) shows that, from a purely national perspective, the optimal combination of policies is a Pigouvian tax on domestic production and a tariff on imports. Even in a cooperative setting, Keen and Kotsogiannis (2014) show that some forms of BCAs are required to achieve global Pareto-efficiency.

In this paper, we focus on the strategic role of BCAs in influencing countries' actions on climate change. In an intra-industry trade model, we are interested to show whether and under which conditions an escalating sequence of BCA-measures, including carbon tariffs and export rebates, could be successful in enforcing a fully cooperative agreement among two countries which differ in their perception of environmental damages.

[1] Output-based rebating is a variant of this and is also sometimes suggested as an anti-leakage measure which implies giving rebates on output produced domestically, regardless whether consumed domestically or exported (Fischer and Fox, 2012).

Our paper contributes to the game-theoretic studies in particular to the strategic environmental-trade policy literature. We solve a three-stage game. In the first stage, we assume that countries decide whether to cooperate on mitigating global emissions based on a multi-stage bargaining game. The environmentally more concerned country takes the initiative, and uses a sequence of escalating threats to enforce cooperation against the environmentally less concerned country. The BCAs threats include not only carbon tariffs, but we also consider complementing carbon tariffs with export rebates. The effect of adding export rebates on the endogenous and strategic decision of countries and on global and individual welfare as well as on global emissions is a prior not evident. Therefore, in this paper, we extend the set up of BCAs on imports, which has been put forward by Eyland and Zaccour (2012; 2014), and add BCAs on exports. We differentiate between two forms of export rebates: optimal and full rebates. Although it is always in the best interest of environmentally concerned governments to impose carbon tariffs, it is not always in their best interest to provide their exporting firms with a full rebate as rebates cost tax payers money. In the second stage, governments choose their climate policy levels under different tax regimes, and, in the third stage, firms choose their equilibrium output levels.

We start our analysis with an initial situation in which two governments impose non-cooperatively a production-based carbon tax on their producers (PB-regime). We first show that only if the individual evaluation of environmental damages in the two countries is similar will both countries be better off under full cooperation (FC-regime). If full cooperation cannot be achieved, we ask whether the environmentally more concerned country can enforce the FC-regime through a sequence of threats leading gradually to a unilateral consumption-based carbon tax. In an escalating penalty game, we consider three threats which constitute various forms of BCAs: (1) a carbon tariff which fully adjusts the difference between the two national tax levels (BI-threat); (2) in addition to carbon tariffs, export rebates where the rebate rate is chosen optimally and hence may not be a full rebate (BIE-threat); (3) carbon tariffs are combined with a full export rebate which implies de facto a unilateral consumption-based tax (BF-threat).

We show under which conditions a weaker punishment is sufficient to establish cooperation and under which conditions penalties need to be escalated. Moreover, we derive conditions such that penalties are credible for the country which tries to enforce cooperation. We show that BCAs can be global welfare distorting, but only under those conditions when they help to enforce cooperation. The harsher a BCA-threat and hence the more effective a BCA-threat is to enforce cooperation, the more distortionary it would be if implemented compared to the FC-regime.

We confirm the "paradox of cooperation" as derived already by Barrett (1994) in our context with trade and BCAs, namely that whenever the global gains from cooperation are expected to be rather significant, BCAs are not sufficient to enforce cooperation.

A wide strand of literature focuses on assessing the effectiveness of BCA-measures in mitigating carbon leakage and the competitive loss of EITE industries operating in countries which impose a carbon price (Babiker and Rutherford, 2005; Böhringer et al., 2012a; Branger and Quirion, 2014; Fischer and Fox, 2012; Winchester, 2011).[2] Most of this literature employs partial or computable general equilibrium (CGE) climate models and concludes that BCAs can effectively mitigate carbon leakage and reduce the output loss of EITE industries, yet they have a small effect on reducing the global costs of mitigation. However, they provide mixed evidence about the importance of adding export rebates to carbon tariffs (Böhringer et al., 2012a; 2014; Branger and Quirion, 2014). Because of assuming unilateral climate policies, these studies do not capture the impact of BCAs in a strategic context. In addition, they assume an exogenous carbon price and hence do not show how climate policy levels would change with BCAs.

This paper is related to three strands of literature, studying the strategic effects of BCAs on the decisions of countries towards mitigating climate change. Broadly speaking, analysing the impact of introducing BCA-measures can be modelled in a two-stage game.[3] In the first stage, countries decide whether to join or comply with a climate agreement to reduce greenhouse gas emissions, and in the second stage they decide on their carbon policy level which we can typically think of as a carbon tax.

The first strand of the literature focuses on the first stage of the game. This strand analyses whether the threat of imposing BCAs is effective to enforce cooperation or compliance with a climate agreement using CGE models (Böhringer et al., 2016; Irfanoglu et al., 2015; Weitzel et al., 2012). For instance, Weitzel et al. (2012), assuming a model of nine regions, analyse whether full BCAs (comprising both carbon tariffs and export rebates) would allow a grand coalition which includes all regions to be stable. They find that BCAs may not be sufficient to induce countries such as China to cooperate. In contrast, Böhringer et al. (2016) consider only carbon tariffs and show that China and Russia respond by cooperation.

[2] There are certain studies that focus on the legal issues of designing and implementing BCAs. See for instance Cosbey et al. (2019) for a recent survey on this issue.

[3] Certain studies, including our paper, consider a three-stage game, where firms choose their production levels in the third stage.

Similarly, though considering three regions only, Irfanoglu et al. (2015) show that imposing carbon tariffs by the United States could be a credible and effective threat to enforce compliance of China with a climate agreement. Due to the complicated nature of the CGE models, these studies consider an endogenous choice of strategies only in the first stage of the game. However, the carbon tax in those countries imposing BCAs is assumed to be exogenous in the second stage. In addition, the optimal response by countries which face BCAs in terms of their policy choice is not modelled. That is, if countries decide not to join the agreement, they remain inactive. Ignoring the optimal choice of carbon taxes undermines some strategic aspects of BCAs.

In contrast to the first strand, the studies in the second strand of the literature focus only on the second stage of the game, i.e. on the endogenous choice of carbon taxes (Eyland and Zaccour, 2012; 2014; Hecht and Peters, 2018; Sanctuary, 2018). However, those studies assume a non-cooperative outcome in the first stage. Most of the studies in this strand are based on theoretical two-country partial equilibrium models. Both Sanctuary (2018), assuming perfect competition, and Eyland and Zaccour (2014), assuming imperfect competition, consider only BCAs on imports which have shown to be effective in pushing both countries to adjust their carbon tax upward. In contrast, Hecht and Peters (2018), assuming also imperfect competition, find that BCA-measures allow the country which imposes these measures to set a higher carbon tax, while the optimal response of the other country is to adjust its tax downward. Their results hold for both BCAs on imports and symmetric BCAs on imports and exports.[4] Therefore, the results related to the policy level of the country which faces BCAs is not clear-cut. Although the political game-theoretic analysis of Helm et al. (2012) abstracts from explicit micro-foundation, Helm et al. show that exporting countries on which carbon tariffs are imposed would respond by taxing their exports rather than remaining inactive or retaliating.

Our paper is closely related to the third strand of the literature which allows the endogenous choice of strategies in both the first and the second stage of the game (Al Khourdajie and Finus, 2018; Anouliés, 2015; Baksi and Chaudhuri, 2017). These studies are based on theoretical models, in particular on strategic imperfect-competition trade model. Al Khourdajie and Finus (2018) show in n-symmetric coalition model that BCAs on imports is an effective measure to reach larger stable climate agreements.[5] Similarly, both Anouliés (2015) and Baksi and

[4] They assume that both measures are imposed at the same rate.

[5] Barrett (1997) also shows in n-symmetric country model that trade bans could lead to full

Chaudhuri (2017) examine the role of BCAs on imports to enforce cooperation, though between two asymmetric countries. Anouliés (2015) assumes that carbon tariffs can be imposed only by the country which unilaterally complies with the cooperative target. She finds that both countries defect from the agreement without BCAs, while carbon tariffs guarantee the compliance of at least one country. For an equal distribution of population between countries, she concludes that BCAs may lead to full cooperation only if countries are nearly symmetric in their environmental damage evaluation. In an infinitely repeated game, Baksi and Chaudhuri (2017) analyse the sustainability of cooperation between the North and South using trigger strategies. They conclude that if the North imposes carbon tariffs only if the South defects from the agreement, BCAs will enhance cooperation if countries are not highly asymmetric in their marginal damages. However, if the North defects and is allowed to implement carbon tariffs, BCAs will hinder cooperation that could be achieved without BCAs. This suggests that BCAs should be implemented only if cooperation cannot be established, a point which we consider in our analysis.

The remainder of the paper is organised as follows. In Section 2, we present the model. In Section 3, we solve for third stage, in Section 4 for the second stage and in Section 5 for the first stage of our three-stage game. In each stage, we derive positive and normative conclusions. Finally, Section 6 concludes and discusses possible future research.

2 Model

We consider two countries, respectively, their governments, $i = 1, 2$, which interact in a strategic trade model. There are two firms, $k = 1, 2$, producing a homogeneous emission-intensive good x which generates greenhouse gas emissions. Firm 1 is located in country 1, and firm 2 is located in country 2. The two firms compete in outputs, i.e., in a Cournot-fashion, and each firm supplies the home and the foreign market. The inverse demand function in market i is given by:

$$p_i(X_i) = a - X_i\,, \quad \forall\ i = 1, 2\,, \tag{1}$$

where p_i is the market price in market i and parameter $a > 0$ is the chock-off price. Total consumption in country i is $X_i = x_{1i} + x_{2i}$, where x_{1i} and x_{2i} are the outputs supplied by firm 1 and 2 to market i, respectively.

cooperation, though trade bans are very different from BCAs.

We solve a three-stage game by backward induction.

In the first stage, countries decide whether to implement a cooperative or non-cooperative climate policy. The cooperative policy corresponds to the implementation of a socially optimal emission tax. The non-cooperative policy comprises four possible regimes: a nationally implemented production-based tax as well as three different designs of border carbon adjustment regimes (BCA-regimes). This stage is modelled as a sequential bargaining game. Country 1, which is the environmentally more concerned country in our model, proposes a cooperative agreement. It uses a sequence of escalating penalties in order to force country 2 to accept its proposal. Those penalties correspond to various forms of unilateral BCA-policies. The details are provided in Section 5.

In the second stage, governments simultaneously choose their climate policy levels cooperatively or non-cooperatively under the different regimes. The analysis of this stage comprises the slopes of best reply functions in the two countries, the ranking of equilibrium tax levels as well as the comparison across regimes regarding global and individual welfare levels.

In the third stage, firms simultaneously choose their output levels for both markets. Those outputs are a function of taxes in both countries, which differ across regimes. The analysis of this stage comprises the relation between taxes and output for the various markets, the relation of producer and consumer surplus in the two countries as well as the effect of taxes on output and hence on global emissions.

In the remainder of this section, we briefly derive equilibrium output of firms in the third stage as a function of what we call "effective taxes". Then we explain the various components of countries' welfare functions, relate those to the different policy regimes and explain what it means for effective and equilibrium taxes.

The outcome of the third stage is a Nash equilibrium in output levels in each of the two markets. We assume completely identical firms with a linear production cost function, i.e., $C_{ki}(x_{ki}) = cx_{ki}$ with $k = 1, 2$ indicating the location of firm k and $i = 1, 2$ the market for which the good is produced. Markets are segmented. That is, firms make separate quantity decisions for both the home and the foreign market. The profits of firms obtained in market 1 and market 2 are given by:

$$Market\,1: \pi_{11} = (p_1(X_1) - c - t_{11})x_{11} \;\&\; \pi_{21} = (p_1(X_1) - c - t_{21})x_{21}\,, \qquad (2)$$

$$Market\,2: \pi_{12} = (p_2(X_2) - c - t_{12})x_{12} \;\&\; \pi_{22} = (p_2(X_2) - c - t_{22})x_{22}\,, \qquad (3)$$

where t_{11} (t_{21}) is the effective carbon tax which firm 1 (2) faces on its supply to market 1 and t_{12} (t_{22}) is the effective carbon tax which firm 1 (2) faces on its supply to market 2; $X_1 = x_{11} + x_{21}$ and $X_2 = x_{12} + x_{22}$. We assume a constant emission-output ratio across firms, which we normalise to 1, such that an emission tax is de facto an output tax.

The simultaneous maximisation of profits obtained in market 1 in (2) and market 2 in (3) by both firms gives the Nash equilibrium quantities supplied by firm 1 and 2:[6]

$$Market\ 1:\ x_{11}^{*} = \frac{A - 2t_{11} + t_{21}}{3}\ \&\ x_{21}^{*} = \frac{A - 2t_{21} + t_{11}}{3}\,, \tag{4}$$

$$Market\ 2:\ x_{12}^{*} = \frac{A - 2t_{12} + t_{22}}{3}\ \&\ x_{22}^{*} = \frac{A - 2t_{22} + t_{12}}{3}\,, \tag{5}$$

with $A = a - c$, which we interpret as a market size parameter. Accordingly, profits of each firm k are given by the sum of profits obtained in market 1 and market 2 as follows:

$$\Pi_1^{*} = \pi_{11}^{*} + \pi_{12}^{*} = (x_{11}^{*})^2 + (x_{12}^{*})^2\ \&\ \Pi_2^{*} = \pi_{21}^{*} + \pi_{22}^{*} = (x_{21}^{*})^2 + (x_{22}^{*})^2\quad. \tag{6}$$

In the second stage, governments choose simultaneously the level of their carbon tax t_i based on the following welfare functions:

$$W_1 = CS_1 + PS_1 + TR_1 - D_1(e) + BCAI_1 - BCAE_1\,, \tag{7}$$

$$W_2 = CS_2 + PS_2 + TR_2 - D_2(e)\,, \tag{8}$$

where CS_i is the consumer surplus in country i, with the consumer surplus being given by $CS_i = \frac{X_i^2}{2}$ which follows from (1), recalling that the total supply to market i is given by $X_i = x_{1i} + x_{2i}$. PS_i is the producer surplus, which is the total profits of firm k based in country i, i.e., $PS_i = \Pi_i^{*}$, as given in equation (6). TR_i is the tax revenue of government i imposed on the production in its country, where $TR_i = t_i\,(x_{k1} + x_{k2})$.

D_i are individual damages from global greenhouse gas emissions released in the production of good x. Global damages from emissions are $D\,(e)$, where $e = X_1 + X_2$ as we normalise the emission-output coefficient to 1. Hence, global emissions are

[6] The Nash equilibrium output levels always exist and are unique in each market. See Appendix A. Moreover, output levels need to be positive for equilibrium tax levels under the various regimes. The sufficient conditions are derived in Appendix A.

equal to total production, which is equal to total consumption. That is,

$$D(e) = de, \quad D_1 = \gamma D(e), \quad D_2 = (1-\gamma)D(e), \gamma \in [0.5, 1] \ , \tag{9}$$

with $d > 0$ a damage parameter, reflecting global marginal damages. We allow for the possibility that countries perceive or evaluate global damages from emissions differently. We assume $\gamma \in [0.5, 1]$. That is, country 1 is at least as concerned as country 2 about environmental damages and usually more whenever γ is strictly larger than 0.5.

The other terms in the welfare function of country 1 in (7) are introduced in the course of the subsequent discussion of the different policy regimes as they are only relevant under the three non-cooperative BCA-regimes. The implications for effective and equilibrium taxes are illustrated in Table 1.

If countries fully cooperate (FC-regime), they maximise $W_1 + W_2$ with respect to a uniform tax t, ignoring the $BCAI_1$ and $BCAE_1$ term in (7). All effective taxes are the same and the equilibrium tax is t^{FC*}, as shown in Table 1.

The cooperative regime is contrasted with four non-cooperative regimes in this paper.

First, in the production-based regime (PB-regime), each government imposes a carbon tax on the production of its home firm. Hence, the effective tax which each firm faces in both markets is equal to the tax imposed by its home country (see Table 1). We denote the corresponding equilibrium taxes by t_1^{PB*} and t_2^{PB*}. As we will show later, $t_1^{PB*} > t_2^{PB*}$ if $\gamma > 0.5$, i.e., the environmentally more concerned country 1 imposes a higher production-based tax.

Table 1: Effective and Equilibrium Carbon Taxes under Cooperative and Non-Cooperative Regimes

	Effective Taxes	FC	PB	BI	BIE	BF
Market 1	t_{11}	t	t_1	t_1	t_1	t_1
	t_{21}	t	t_2	t_1	t_1	t_1
Market 2	t_{12}	t	t_1	t_1	$t_1(1-\varphi^*)$	0
	t_{22}	t	t_2	t_2	t_2	t_2
Equilibrium Taxes $\forall\, i = 1, 2$		$t = t^{FC*}$	$t_i = t_i^{PB*}$	$t_i = t_i^{BI*}$	$t_i = t_i^{BIE*}$	$t_i = t_1^{BF*}$

Second, in the border carbon adjustment on imports regime (BI-regime), additionally to a production-based tax, country 1 imposes a tariff on imports from country 2. Thus, firm 1 faces the effective tax $t_1 = t_{11} = t_{12}$ as under the PB-regime, and also firm 2 faces $t_{22} = t_2$ on its supply to country 2 as above but the effective tax t_{21} on its export to country 1 is now given by $t_{21} = t_2 + \omega(t_1 - t_2)$.

Country 1 is only allowed to impose the tariff provided $t_1 > t_2$ with ω the border tax adjustment parameter on imports (Eyland and Zaccour, 2012). That is, under the BI-regime, the term $BCAI_1$ is included in country 1's welfare function with $BCAI_1 = \omega\,(t_1 - t_2)\,(x_{21})$ if $t_1 > t_2$, otherwise $BCAI_1 = 0$. We assume henceforth $\omega = 1$ for two reasons. Any value of ω above 1 would be illegal under the rules of the World Trade Organization (WTO).[7] Additionally, given the opportunity of using carbon tariffs, any value smaller than 1 would not be optimal for country 1.[8] This assumption implies that both firms supplying market 1 face the same effective carbon tax. (see Table 1). Therefore, BCAs on imports always constitutes "full adjustment" on imports. Equilibrium taxes, which follow from maximisation of each government's individual welfare function with respect to own taxes, gives t_1^{BI*} and t_2^{BI*}.

Third, in the border carbon adjustments on imports and exports regime (BIE-regime), country 1 complements its production-based tax and carbon tariff on imports with a rebate on its firm's exports to country 2. We assume that the export adjustment/rebate rate φ is chosen optimally by the government. Compared to the BI-regime, the effective tax on exports of firm 1 to country 2, t_{12}, is now given by $t_{12} = t_1(1-\varphi)$, again, only if $t_1 > t_2$, with φ the border tax adjustment parameter on exports, or, equivalently the export rebate rate.[9] According to the WTO-rule of non-discrimination, $t_1(1-\varphi) \geq t_2$ must hold. Generally speaking, the optimal φ, which we denote by φ^* can be positive or negative and can be smaller or larger than 1. Later we clarify the range of φ^*. At this stage it suffices to point out that it is neither obvious that φ^* is chosen such that $t_1(1-\varphi^*) = 0$ (full rebate) nor such that $t_1(1-\varphi^*) = t_2$ (full adjustment) because subsidising exports is costly. Under the BIE-regime, the term $BCAE_1$ is included in country 1's welfare function with $BCAE_1 = \varphi\, t_1\, x_{12}$ if $t_1 > t_2$ and $t_1(1-\varphi) \geq t_2$, otherwise $BCAE_1 = 0$. Equilibrium taxes follow from the individual maximisation of welfare functions with respect to own taxes, leading to t_1^{BIE*} and t_2^{BIE*}.

Fourth, the border carbon adjustments on imports with full export rebate regime (BF-regime) is similar to the BIE-regime, except that the rebate rate on exports φ is not chosen optimally but set to one, i.e., φ= 1. Hence, firm 1 faces effective tax $t_{12} = 0$ on its exports. Equilibrium taxes are given by t_1^{BF*} and t_2^{BF*}.

[7]The General Agreement on Tariffs and Trade (GATT) allows WTO members to apply a border tax adjustment at a rate, which is not higher than the rate applied to domestically produced "like" products.

[8]See Hecht and Peters (2018) and Weitzel et al. (2012). That is, if ω could be chosen endogenously under the restriction $\omega \leq 1$, $\omega^* = 1$.

[9]Note that we cannot model BCAs on exports in the same way as on imports because country 2 does not tax firm 1. That is, we cannot assume for instance $t_{12} = t_1 - \varphi(t_1 - t_2)$.

3 Third Stage

In this section, we have a closer look at the effect of taxes on equilibrium output levels. These effects are straightforward in the fully cooperative regime with a uniform tax. All output levels for each market are the same. Increasing taxes gradually lowers all outputs uniformly, and decreases profits and consumer surplus, also all uniformly. The benefits are uniformly higher tax revenues and lower damages, even though, country 1 benefits more than country 2 from lower damages as long as $\gamma > 0.5$.

In the non-cooperative policy regimes, taxes in country 1 and 2 will be different, with $t_1 \geq t_2$ and strict inequality if $\gamma > 0.5$. Also, the effects of taxes will be different from those under the cooperative regime.

Proposition 1. The Effect of Non-Cooperative Taxes on Equilibrium Production and Consumption

i. The production of firm 1 (firm 2) for market 1:

- *decreases (increases) in the tax of country 1, while increases (decreases) in the tax of country 2 under the PB-regime;*
- *decreases (decreases) in the tax of country 1 and is independent (independent) of the tax of country 2 under all BCA-regimes.*

ii. The production of firm 1 (firm 2) for market 2:

- *increases (decreases) in the tax of country 2 under all regimes;*
- *decreases (increases) in the tax of country 1 under the PB- and the BI-regime, and under the BIE-regime if $\varphi^* < 1$;*
- *increases (decreases) in the tax of country 1 under the BIE-regime if $\varphi^* > 1$;*
- *is independent (independent) of the tax of country 1 under the BF-regime.*

iii. The consumption in both markets:

- *decreases in the tax of both countries under the PB-regime;*
- *Under all BCA-regimes:*
  - *the consumption in market 1 decreases in the tax of country 1, while it is independent of the tax of country 2;*

- *the consumption in market 2 always decreases in the carbon tax of country 2, decreases (increases) in the tax of country 1 if* $\varphi^* < 1$ $(\varphi^* > 1)$ *under the BIE-regime, and is independent of the tax of country 1 under the BF-regime.*

Proof. Follows from inserting the effective taxes in Table 1 into (4) and (5) and differentiating output with respect to tax levels. □

Under the PB-regime, the standard effects as known from the literature hold: the production of each firm decreases in its domestic tax level while it increases in the foreign tax level. Consumption levels in both markets are negatively affected by both taxes. Therefore, facing a higher carbon tax, firm 1 is less competitive in both markets, though the consumer surplus is the same in both countries.

With an import tariff, which is part of all three BCA-regimes (BI-, BIE- and BF-regime), country 1 fully controls the output supplied to its country. Taxes on the supply to country 1 are fully adjusted through import tariffs. Hence, all outputs produced for market 1 (and hence also consumption in country 1) are negatively affected by tax t_1 and are independent of tax t_2. Thus, profits in market 1 are the same for both firms, though profits in market 2 are still lower for firm 1 than firm 2 (because $t_1 > t_2$ if $\gamma > 0.5$). It is also clear that consumers in country 1 enjoy a lower consumer surplus because the supply to market 1 is lower than to country 2.

If a carbon tariff is supplemented by an export rebate under the BIE- and BF-regime, the output of firm 1 (firm 2) sold to market 2 is still negatively (positively) affected by the tax in country 1 if $\varphi^* < 1$ (which is one possibility under the BIE-regime) though to a lesser extent than under the PB- and BI-regime and is unaffected by the tax in country 1 if $\varphi = 1$ as under the BF-regime and even increases (decreases) in t_1 if $\varphi^* > 1$ (which is another possibility under the BIE-regime).[10] The same relations are found for total consumption in market 2. However, regardless of the value of φ, firm 1's profits will be (weakly) lower than firm 2's profits in market 2 because we always have: $t_1(1-\varphi) \geq t_2$. Thus, export rebates will usually not level the playing field in market 2, though it lowers the difference in profits between the two firms. Accordingly, also the difference in consumer surpluses in the two countries as observed under the BI-regime will

[10]Overcompensating firms is not uncommon. For example, Martin et al. (2014) show that the free allocation of emission permits under the European Union Emissions Trading Scheme (EU-ETS) resulted in sizeable overcompensation of emission-intensive industries.

remain, and in fact may be become even larger as exports to market 2 are subsidised under the BIE- and BF-regime. Interestingly, under the BF-regime which is de facto a unilateral consumption-based tax, the consumption in each market is independent of the foreign country tax level.

Thus, moving from the PB-regime to the BI-regime, raises profits of firm 1 in market 1, but disadvantages consumers in country 1. Moving from the BI-regime to the two regimes with export rebates (BIE- and BF-regime), improves upon the profits of firm 1 in market 2, but drives a further wedge between the consumer surplus enjoyed in market 1 and 2. Of course, in order to understand how the different regimes impact on the welfare levels of both countries, equilibrium tax levels need to be considered, which is part of the second stage of the three stage game and treated in Section 4. Moreover, taxes, tariffs and rebates do not only affect producers and consumers, but also affect revenues of governments. Finally, one motivation of BCAs is the possibility of reducing environmental damages. On the way of clarifying this issue, we offer Proposition 2.

Proposition 2. The Effect of Non-Cooperative Carbon Taxes on Global Emissions

- $\frac{\partial e^{PB}}{\partial t_1} = \frac{\partial e^{PB}}{\partial t_2} < 0.$
- *The carbon tax of country 1 has the largest impact on reducing global emissions under the BI-regime if* $\varphi^* > 0$*:*
  i. $\left|\frac{\partial e^{PB}}{\partial t_1}\right| = \left|\frac{\partial e^{BF}}{\partial t_1}\right| \leq \left|\frac{\partial e^{BIE}}{\partial t_1}\right| < \left|\frac{\partial e^{BI}}{\partial t_1}\right| < 0$ *if* $\varphi^* \leq 1$,
  ii. $\left|\frac{\partial e^{BIE}}{\partial t_1}\right| < \left|\frac{\partial e^{PB}}{\partial t_1}\right| = \left|\frac{\partial e^{BF}}{\partial t_1}\right| < \left|\frac{\partial e^{BI}}{\partial t_1}\right| < 0$ *if* $1 < \varphi^* < 3$*, and*
  iii. $\frac{\partial e^{BIE}}{\partial t_1} \geq 0$ *if* $\varphi^* \geq 3$.
- *The carbon tax of country 2 has a lower impact on reducing global emissions under the BCA-regimes than under the PB-regime:* $\left|\frac{\partial e^{BF}}{\partial t_2}\right| = \left|\frac{\partial e^{BIE}}{\partial t_2}\right| = \left|\frac{\partial e^{BI}}{\partial t_2}\right| < \left|\frac{\partial e^{PB}}{\partial t_2}\right| < 0.$

Proof. Follows from inserting the effective taxes in Table 1 into (4) and (5), and differentiating total output with respect to taxes, recalling that we assume a constant emission-output ratio equal to 1. □

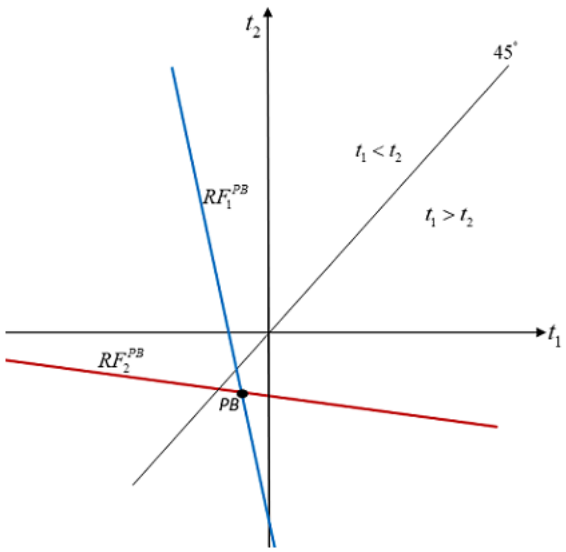
Under the PB-regime, the climate policy of both governments has the same impact on reducing global emissions. However, moving from the PB- to the BI-regime, the climate policy in country 1 has a stronger impact on global emissions as it now controls not only its domestic production but also imports. Adding export rebates may partially (if $\varphi < 1$) or completely (if $\varphi = 1$) offset the effect of carbon tariffs on taxes in country 1 to reduce global emissions. In fact, for $\varphi > 1$, the effect of taxes t_1 in reducing global emissions is even lower than under the PB-regime. Hence, complementing carbon tariffs with export rebates reduces the effectiveness of taxes in country 1 in reducing global emissions. Moreover, BCA-measures generally weakens the impact of climate policy in country 2 on global emissions. Thus, in order to have the full picture how the different regimes impact on producers, consumers, revenues and damages, and hence on the strategic interaction among the two countries including leakage effects, we need to consider the second stage.

4 Second Stage

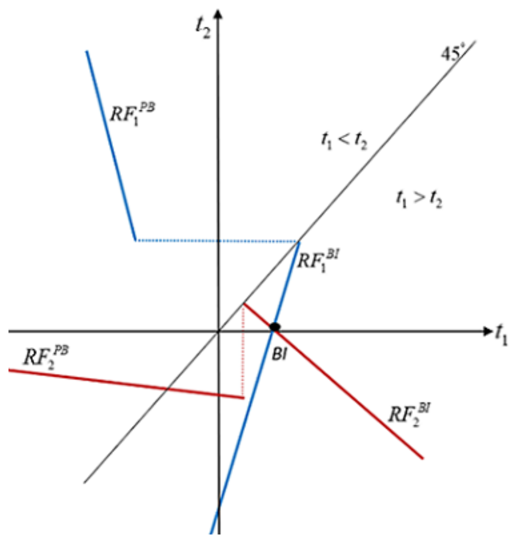
In the second stage, governments choose their climate policy level cooperatively or non-cooperatively. In Appendix A, we derive equilibrium taxes under the full cooperative regime (FC-regime) and the four non-cooperative regimes (PB-, BI-, BIE- and BF-regime). We establish existence and uniqueness of equilibria and also derive what we call non-negativity constraints (NN-constraints) and border carbon adjustment constraints (BCA-constraints) that impose conditions on the parameters of our model. NN-constraints are simply conditions such that all output levels are positive and hence establish interior solutions. Typically, they require β to be larger than some threshold say $\breve{\beta}$ with $\beta = \frac{A}{d}$ and $A = a - c$. We recall that A is a measure of the market size of our model, or, a proxy of the net benefits of production and consumption whereas d is the parameter that evaluates damages. Hence, if d would be too large in relation to A, taxes would imply negative outputs. BCA-constraints ensure that equilibrium taxes in country 1 are higher than in country 2. Such conditions can also be expressed in terms of β to be smaller than some threshold say $\hat{\beta}(\gamma)$. Whenever we conduct a comparison across regimes, we assume the most restrictive NN- and BCA-constraints to hold.

In Appendix A, we also derive reaction functions under the non-cooperative regimes, which are illustrated in Figure 1. We refer to country 1's reaction function as RF_1 and that of country 2 as RF_2. Figure 1 assumes particular parameter values, though the sign of the slope of the reaction functions (though not the slope itself) would be the same for other parameter values.

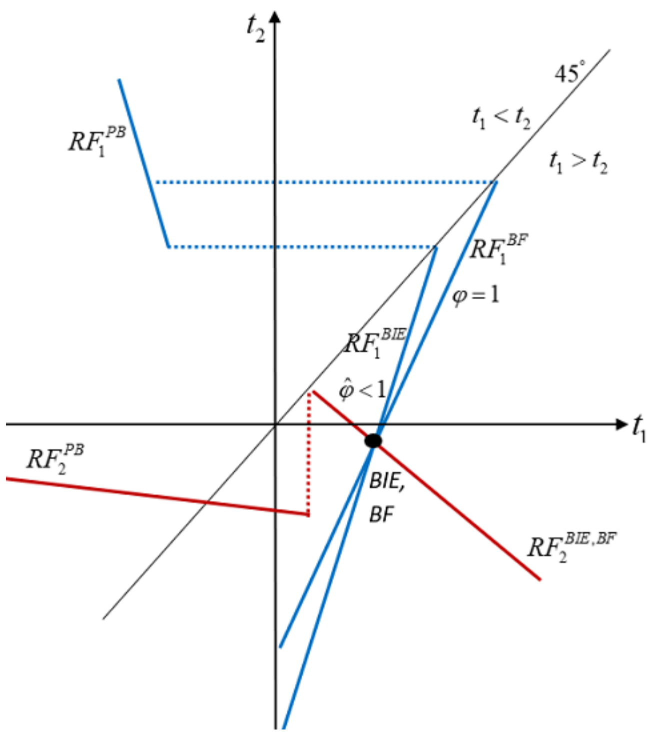
In panel (a) in Figure 1, reaction functions under the PB-regime are shown. As expected, reaction functions are downward sloping implying that taxes are strategic substitutes. In panel (b) and (c) reaction functions under the three BCA-regimes are drawn. The reaction function of country 2, the country on which the



(a) PB-regime



(b) BI-regime



(c) BIE- and BF-regime

Figure 1: Reaction Functions of Countries under Non-cooperative Regimes
*The reaction functions are drawn for: $A = 10$, $d = 3$, and $\gamma = 0.7$.

BCA-measures are unilaterally imposed, remains downward sloping. In contrast, country 1' reaction function becomes upward sloping for all $t_1 > t_2$, which is the region below the 45°-line.

Under all BCA-regimes, not only country 1's but also country 2's reaction function is a piecewise reaction function with a jump at the 45°-line. For country 1, with $RF_1 := t_1(t_2)$, increasing t_2 gradually, there is a level of $\bar{t_2}$ for which matching taxes, i.e., $t_1(\bar{t_2}) = \bar{t_2}$, is a best response. At this point, and for any higher tax t_2, the reaction function jumps from the BCA-regime to the PB-regime as now $t_1(t_2) \leq t_2$. A similar explanations applies to the reaction function of country 2, $RF_2 := t_2(t_1)$ (see details in Appendix A.7).

Adding tariffs to the PB-regime, as done under the BI-regime, implies that carbon taxes become strategic complements for country 1. The intuition is as follows. First, unlike under the PB-regime, a higher level of t_2 has no effect on the consumption in country 1. Second, the impact of t_2 on global emissions become smaller compared to the PB-regime (and the impact of t_1 larger; see Proposition 2). Hence, a higher level of t_2 is not sufficient to reduce damages in country 1 and hence country 1 responds by raising its tax. Third, carbon tariffs revenues create a new incentive for country 1 to tax emissions. Recalling that tariff revenues depend on the difference between the two tax levels, country 1 raises its tax level if t_2 increases in order to obtain higher tariff revenues.

The reaction function of country 2, RF_2, remains downward sloping, though it becomes steeper under the BI-regime for t_1 sufficiently large, i.e. below the 45°-line. For a given t_1, country 2 has an incentive to raise its carbon tax level in order to protect its tax revenues captured by country 1. However, an increase in t_1 also harms consumers in country 2. Moreover, it reduces the profits of firm 2 obtained in market 1 as firm 2 faces t_1 on its exports. The overall effect is that country 2 reduces its tax more strongly than under the PB-regime for any marginal increase of t_1.

The slope of country 1's reaction function RF_1 is also upward sloping after complementing carbon tariffs with export rebates under the BIE-and BF-regime. However, after adding export rebates, RF_1 becomes flatter in panel (c) than in panel (b) and the slope decreases as φ increases in the t_2-t_1-space if φ is not sufficiently high (i.e., $\partial RF_1/\partial t_2$ becomes larger and increases in φ).[11] For a marginal increase of t_2, country 1 has a larger incentive to raise its tax level with export rebates which allows country 1 to address the competitive loss of its firm in market 2.

[11] See Appendix A.7.

Moreover, facing less carbon leakage raises the effectiveness of country 1's carbon tax, aiming at internalising its damages from global emissions. For country 2's reaction function, export rebates make no difference. The intersection of RF_2 with the flatter RF_1 leads to a higher tax level in country 1 under the BIE- and BF-regime compared to the BI-regime. However, because taxes are strategic substitutes for country 2, this country responds with a less stringent climate policy. A higher level of t_1 means lower profits of country 2' firm obtained in market 1. Additionally, export rebates reduce the market share of firm 2 in its home market in country 2. Hence, country 2 sets a lower tax t_2 to protect the competitiveness of its home firm in market 2.

We conclude the above discussion with Proposition 3, which ranks the Nash equilibrium climate policy levels across the different regimes.

Proposition 3. Ranking of Equilibrium Carbon Taxes

- *Equilibrium taxes in country 1 can be ranked as follows:* $t_1^{PB*} < t_1^{BI*} < t_1^{BIE*} = t_1^{BF*}$.
- *Equilibrium taxes in country 2 can be ranked as follows:* $t_2^{PB*} < t_2^{BIE*} = t_2^{BF*} < t_2^{BI*}$.
- *Under the BIE-regime,* $\varphi^* > 0$, *and* $\varphi^* \leq (>) 1$ *if* $\beta \leq (>) \ddot{\beta}(\gamma)$ *with* $\ddot{\beta}(\gamma)$ *increasing in* γ *and* $\beta = \frac{A}{d}$.
- *The effective taxes of country 1 in market 1 are the equilibrium taxes in country 1. Hence,* $t_{11}^{PB*} < t_{11}^{BI*} < t_{11}^{BIE*} = t_{11}^{BF*}$.

  The effective taxes of country 1 in market 2 can be ranked, depending on the values of β *and* γ*, as follows: i.* $t_{12}^{PB*} < t_{12}^{BF*}, t_{12}^{BIE*} < t_{12}^{BI*}$*, ii.* $t_{12}^{BF*} \leq (>) t_{12}^{BIE*}$ *if* $\varphi^* \leq (>) 1$*, and iii.* $t_{12}^{BF*} < t_{12}^{PB*} < t_{12}^{BIE*} < t_{12}^{BI*}$.
- *The effective taxes of country 2 in market 2 are the equilibrium taxes in country 2. Hence,* $t_{22}^{PB*} < t_{22}^{BIE*} = t_{22}^{BF*} < t_{22}^{BI*}$.

  Except for the PB-regime, the effective taxes of country 2 in market 1 are the equilibrium taxes of country 1. Thus, $t_{21}^{PB*} < t_{21}^{BI*} < t_{21}^{BIE*} = t_{21}^{BF*}$.
- *The socially optimal tax level is always larger than the non-cooperative taxes under the PB-regime:* $t_i^{PB*} < t^{FC*}\ \forall\ i = 1, 2$. *However, under the BCA-regimes,* $t^{FC*} < t_i^{BI,BIE,BF*}$ *is possible for one or both countries if global marginal damages* d *are relatively small to the market size* A*, and countries are sufficiently asymmetric, i.e.,* γ *is sufficiently large.*

Proof. See Appendix A.8. □

Proposition 3 makes a distinction between equilibrium taxes and effective taxes.[12] The reason is that firm 2 faces de facto not tax t_2 but tax t_1 on all exports to market 1 under all BCA-regimes. Moreover, firm 1 faces de facto not t_1 but $t_1(1-\varphi)$ on all exports to market 2 under the BIE- and BF-regime with $\varphi = \varphi^* > 0$ as we learn from Proposition 3 under the BIE-regime and $\varphi = 1$ by assumption in the BF-regime. That is, the effective tax of country 1 levied on its firm on exports is lower than the equilibrium tax.

Proposition 3 clearly shows that country 1 chooses higher equilibrium taxes if BCA-measures are available. BCA-measures improve the competitiveness of its firm. Moreover, it controls a larger share of global emissions. That is, country 1 can better address carbon leakage, and, hence, country 1 is more effective in internalising its environmental damages. Additionally, tariffs provide a source for revenues.

The effective tax of country 1 in market 1 is higher under the BIE- and BF-regime than under the BI-regime, but in market 2 this is reversed. Thus, country 1 compensates higher effective taxes in market 1 with lower effective taxes in market 2 if export rebates are available. Since the effective tax of country 1 in market 2 is zero under the BF-regime, i.e., $t_{12}^{BF*} = 0$, it is possible that it is even lower than under the PB-regime, $t_{12}^{BF*} < t_{12}^{PB*}$.

The ranking of the effective taxes of country 1 in market 2 under the BIE- and BF-regime simply depends on whether φ^* is larger or smaller than 1 under the BIE-regime, where we recall that $\varphi = 1$ under the BF-regime. As Proposition 3 suggests, if damages are sufficiently large compared to the net benefits from production and consumption, i.e., β is sufficiently small ($\beta < \ddot{\beta}(\gamma)$), φ^* will be smaller than 1. Apart from balancing net benefits and damages when choosing φ^*, export rebates are costly to country 1.

Country 2 also chooses higher taxes under the BCA-regimes, suggesting that BCAs help to reduce global emissions. Of course, the effective taxes in market 1 are those of country 1, but even effective taxes in market 2 are higher under the BCA-regimes than under the PB-regime. Export rebates reduce the push on country 2 to adjust its taxes upward compared to import tariffs only, as country 2 tries to protect the competitiveness of its firm in its home market.

[12] The fact that equilibrium taxes under the BIE- and BF-regime are equal is a result of our assumption of a linear damage function.

Finally, full cooperation among countries always leads to setting a higher carbon tax than under no cooperation without BCAs. However, since BCAs create incentives for both countries towards setting higher taxes, equilibrium taxes may even be higher than under full cooperation. Important for the understanding of this result is that the socially optimal tax is not equal to global marginal damages due to imperfect competition (this result is well-known in the literature, see, e.g. Barnett (1980); Kennedy (1994) and Duval and Hamilton (2002)). A sufficient condition for both governments to set higher taxes under the BCA-regimes than in the social optimum is when global marginal damages d are low compared to the net benefits from production and consumption A and countries are sufficiently asymmetric, i.e., γ is sufficiently large.[13] In such cases, socially optimal taxes will be low but BCAs still provide an opportunity to protect markets, increase tariff revenues and protect tax revenues.

It is obvious from the above results that global emission levels generally decrease with BCAs compared to the PB-regime. However, the effect of adding export rebates to carbon tariffs is a priori ambiguous. The following Proposition ranks the global emission levels across the BCA-regimes.

Proposition 4. Ranking of Equilibrium Global Emissions

- *BCA-regimes vs PB-regime:* $e^{BI,BIE,BF*} < e^{PB*}$.
- *Across the BCA-regimes:*
  - $e^{BI*} < e^{BIE*}$ *and* $e^{BI*} \gtrless e^{BF*}$;
  - $e^{BIE*} \leq (>) e^{BF*}$ *if* $\varphi^* \leq (>) 1$.

Proof. See Appendix A.9. □

Despite export rebates support the climate policy of country 1, larger supply by firm 1 in market 2 leads to higher global emissions under the BIE-regime compared to the BI-regime. This result confirms the main argument against export rebates. They are less effective in reducing global emissions. An exception is when the relative importance of global marginal damage d is low compared to the net benefits from production and consumption A and if countries are sufficiently asymmetric (see Appendix A.9). In such cases, a lower consumption level in country 1, driven by a higher carbon tax under the BF- than BI-regime, compensates

[13] More precisely, it can be shown that if $\beta > (5 + 8\gamma) \forall \gamma > 0.72$, taxes of both countries under all BCA-regimes are higher than in the social optimum. For a qualitatively similar result, see Baksi and Chaudhuri (2017) for the case of carbon tariffs though under different conditions.

for the effect of adding export rebates. As expected, the relation between global emissions under the BIE and BF-regime depends on the optimal value of φ^*.

Although studying the effects of BCAs on global emissions is important, given that these measures have been proposed for environmental reasons, it is also important to investigate the impacts of these measures on welfare. From a normative point of view, it is important because BCAs do not only affect environmental damages, but also production and consumption. Moreover, in order to understand the incentive of countries if implementing a cooperative climate policy, the effects on individual countries also are important. Given that BCAs are unilateral measures, one intuitively expects that effects for the two countries are different, and, may even go in opposite direction (e.g., Babiker and Rutherford 2005, Böhringer et al. 2012b and Mattoo et al. 2009). Hence, if there are winners and losers, BCAs may lead to an ambiguous effect on global welfare.

Proposition 5. Ranking of Equilibrium Global Welfare

Let $W^* = W_1^* + W_2^*$ be equilibrium global welfare and recall $\beta = \frac{A}{d}$ and $A = a - c$.

1) FC-regime vs PB-regime:

- $W_1^{FC*} > W_1^{PB*}$ *and* $W_2^{FC*} \geq (<) W_2^{PB*}$ *if* $\gamma \leq (>) \overline{\gamma} = \frac{41}{64} \backsimeq 0.64$.
  Let $\Delta W = W^{FC*} - W^{PB*}$, *then* ΔW *decreases in* β.

2) BCA-regimes vs PB-regime:

- $W^{BI*} > W^{PB*}$ *except if* $\beta > \underline{\beta}_W^{BI}(\gamma)\ \ \forall \gamma > \gamma_1$.
- $W^{BIE*} > W^{PB*}$ *except if* $\beta > \underline{\beta}_W^{BIE}(\gamma)\ \ \forall \gamma > \gamma_2$.
- $W^{BF*} > W^{PB*}$ *except if* $\beta > \underline{\beta}_W^{BF}(\gamma)\ \ \forall \gamma > \gamma_3$.
  with $\underline{\beta}_W^{BI} > \underline{\beta}_W^{BIE} > \underline{\beta}_W^{BF}$ *and* $\gamma_1 > \gamma_2 > \gamma_3$.
- *Compared to the PB-regime, country 1 is better off under all BCA-regimes and country 2 is better off if γ is sufficiently small and worse off if γ is sufficiently large.*

3) Across the BCA-regimes:

- $W^{BI*} > W^{BIE*}$ *and* $W^{BI*} > W^{BF*}$.
- $W^{BIE*} \geq (<) W^{BF*}$ *if* $\varphi^* \leq (>)\, 1$.

Proof. See Appendix A.10. □

Regarding the first part of Proposition 5, axiomatically, global welfare in the social optimum is strictly larger than under any other regime. Given our assumption of $\gamma \in [0.5, 1]$, it is not surprising that if γ is larger than some threshold, i.e. if the damage evaluation in country 2, $(1 - \gamma)$, is not sufficiently high, country 2 would be worse off under full cooperation compared to the PB-regime. In Section 5, we take this as defining feature why cooperation is difficult to implement, i.e., we assume $\gamma > \overline{\gamma}$ as the starting point of our analysis. At this point, it suffices to demonstrate that moving from a non-cooperative production-based tax to a fully cooperative production-based tax, the potential global gain from cooperation ΔW decreases in β. That is, full cooperation would matter if the damage parameter d is large compared to the net benefits from the production and consumption parameter A.

The second part of Proposition 5 highlights that global welfare generally increases if country 1 imposes BCA-measures compared to the PB-scenario, though there are exceptions. As country 1 is always better off with the implementation of BCA-measures than in the PB-regime, those exceptions occur if country 2 is worse off under the BCA- than PB-regime. Country 2 is always worse off if its damage evaluation is low, i.e., γ is high, as it benefits very little from the reduction of global emissions under the BCA-regimes in the form of reduced damages. For intermediate values of γ, country 2 is also worse off if β is sufficiently large, as the reduction of damages cannot compensate for the loss of the net benefits from production and consumption. Only for very low values of γ and hence countries are nearly symmetric, will also country 2 be better off under the BCA- than PB-regime. (For details, see Appendix A.10). All together, BCAs increase global welfare compared to the PB-regime if β is not too large, i.e., under those conditions when the potential gains from full cooperation are not very small.

The ranking of thresholds $\underline{\beta}_W^{BI} > \underline{\beta}_W^{BIE} > \underline{\beta}_W^{BF}$ and $\gamma_1 > \gamma_2 > \gamma_3$ in the second part and the third part of Proposition 5 suggests that the implementation of the BI-regime causes less global welfare distortions than the two BCA-regimes with export rebates. Moreover, as long as $\varphi^* < 1$, the BIE-regime causes less distortions than the BF-regime with a full rebate.[14] Thus, when enforcing cooperation through a threat of escalating penalties, one should proceed along the sequence BI-, BIE- and BF-regime in order to minimise the distortions in case those threats

[14] Even if $\varphi^* > 1$, we may have a situation where $W^{BIE*} > W^{BF*}$ as shown in Appendix A.10.

need to be implemented, an idea that we take up in Section 5.

5 First Stage

5.1 Preliminaries

In this section, we solve the first stage of the game. We ask the question under which conditions BCAs can enforce full cooperation and under which conditions this is not possible. We derive the equilibrium path of an escalating penalty game in subsection 5.2, and evaluate equilibria from a global welfare perspective in subsection 5.3.

The escalating game we have in mind is a sequential game with multiple stages as shown in Figure 2. In each stage, country 1 moves first and then country 2. In order to render the analysis interesting, we make the following assumption.

Assumption:

Country 2 has no incentive to fully cooperate if the alternative is the PB-regime. That is, $\gamma > \overline{\gamma}$ with $\overline{\gamma}$ as defined in Proposition 5.

That is, in stage 0, the proposal "cooperation" by country 1 is not accepted by country 2. Recall that country 1 is always better off under full cooperation than the PB-regime (see Proposition 5). Hence, the equilibrium path in stage 0 is the bold highlighted branch which directly leads to stage I. That is, the interesting part of the game starts in stage I.

In stage I, the escalating penalty game starts. Country 1 threats with implementing import tariffs under the BI-regime if country 2 does not accept cooperation. If country 2 accepts, the game ends in node 4 in Figure 2. If country 2 declines, the game proceeds to stage II.

In stage II, country 1 can implement its threat, i.e. the BI-regime, and the game ends at node 5, or, can instead escalate the threat which is the threat to implement the BIE-regime with the aim to enforce cooperation. Country 2 can either give in to the BIE-threat and cooperate in which case we end up in node 6 or it refuses and the game proceeds to stage III.

In stage III, country 1 either implements the BIE-threat and the game terminates at node 7 or escalates the threat to the BF-threat. If country 2 gives in, cooperation is established at node 8. If country 2 still refuses, country 1 will implement the BF-threat and the game ends at node 9. Alternatively, country 1 will implement its most preferred BCA-regime, which, as we will show later, will depend on the parameters of the model.

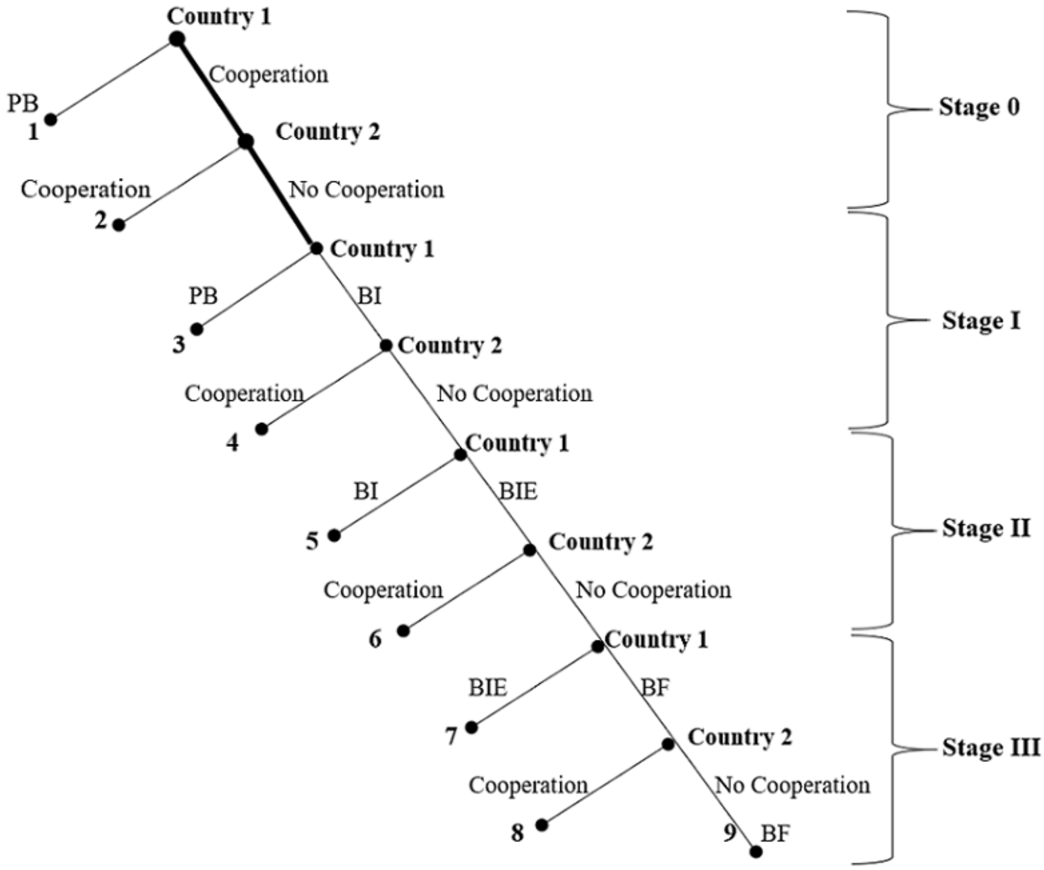


Figure 2: Escalating Penalty Game

As will be evident from the analysis below, there is a unique equilibrium path for each possible parameter range in our model. The sequence depicted in Figure 2 is motivated by our analysis in Section 4, in particular Proposition 5. The enforcement game starts with the least distortionary threat should it be implemented and only escalates the threat should this prove necessary, i.e., country 2 does not accept cooperation. In Appendix B.4, we show that the same equilibrium outcomes would emerge in a one-stage game in which country 1 chooses its action first and country 2 subsequently, though there exist parameter ranges for which the equilibrium is not unique. We solve both games by backward induction in order to derive the subgame-perfect Nash equilibrium.

5.2 The Equilibrium Escalating Penalty Path

For the analysis of the equilibrium path, two features are central. The first feature relates to escalating penalties. We ask the question under which conditions will country 2 accept cooperation for a particular threat and under which conditions will it refuse to cooperate. The second feature is the credibility of threats. We ask the question whether and under which conditions country 1 has an incentive to escalate penalties. The first question is answered in Lemma 1.

Lemma 1. The Effect of BCAs on the Incentive of Country 2 to Cooperate

i. Under the threat of BCAs with tariffs on imports (BI-threat), country 2 is willing to cooperate if $\beta \geq \underline{\beta_1}(\gamma)$.

ii. Under the threat of BCAs with a tariff on imports and an optimal export rebate (BIE-threat), country 2 is willing to cooperate if $\beta \geq \underline{\beta_2}(\gamma)$.

iii. Under the threat of BCAs with a tariff on imports and a full export rebate (BF-threat), country 2 is willing to cooperate if $\beta \geq \underline{\beta_3}(\gamma)$.

We have: $\underline{\beta_1}(\gamma) > \underline{\beta_2}(\gamma) > \underline{\beta_3}(\gamma)$ *with* $\beta = \frac{A}{d}$.

Proof. See Appendix B.1. □

Lemma 1 is illustrated in Figure 3. On the vertical axis we have parameter β and on the horizontal axis parameter γ. The upward sloping straight line 'BCA-C' is the BCA-constraint, implying that only values below this line satisfy the BCA-constraint. The upward sloping straight line 'NN-C', starting at $\gamma \approx \overline{\gamma} = 0.64$, is the non-negativity constraint, implying that only parameter values above this line satisfy this condition.

The black area denoted by PB implies that country 2 would cooperate if faced with the alternative of the PB-regime. We know this would only happen if $\gamma \leq \overline{\gamma}$, which we have ruled out by assumption. The blue area denoted by BI is the parameter range in which the BI-threat enforces cooperation, i.e., condition i holds in Lemma 1. The green area denoted by BIE is the additional parameter space that can enforce cooperation with the BIE-threat. That is, condition ii in Lemma 1 holds in the green and blue area and hence comprises a larger parameter space. The red area is the additional parameter space in which the BF-threat is successful to establish cooperation. Hence, condition iii in Lemma 1 comprises the red, green and blue area. Finally, the grey area is the parameter space in which condition iii in Lemma 1 does not hold. Thus, Lemma 1 confirms the logic that escalation proceeds along the BI-, BIE- and BF-threat.

Unlike the PB-regime, the decision of country 2 to cooperate if faces the BCA-threats depends on both β and γ. BCAs are successful in enforcing cooperation of country 2 if damages are not sufficiently high compared to the net benefits from production and consumption, i.e., if $\beta(\gamma)$ is not sufficiently small. In addition, as is clearly shown in Figure 3, the difference between the three thresholds, $\underline{\beta_1}(\gamma)$, $\underline{\beta_2}(\gamma)$ and $\underline{\beta_3}(\gamma)$ decreases as γ increases, implying that escalating BCA-threats through adding export rebates becomes less effective when countries are highly asymmetric. Taken together, if $\beta(\gamma)$ is sufficiently large, country 2 may accept cooperation if the alternative is any of the BCA-regimes, irrespective of its evaluation of damages. Whereas if $\beta(\gamma)$ is not too large, country 2 is more likely to accept cooperation only if the alternative is the BIE- or BF-regime if its damage evaluation $(1-\gamma)$ is high.

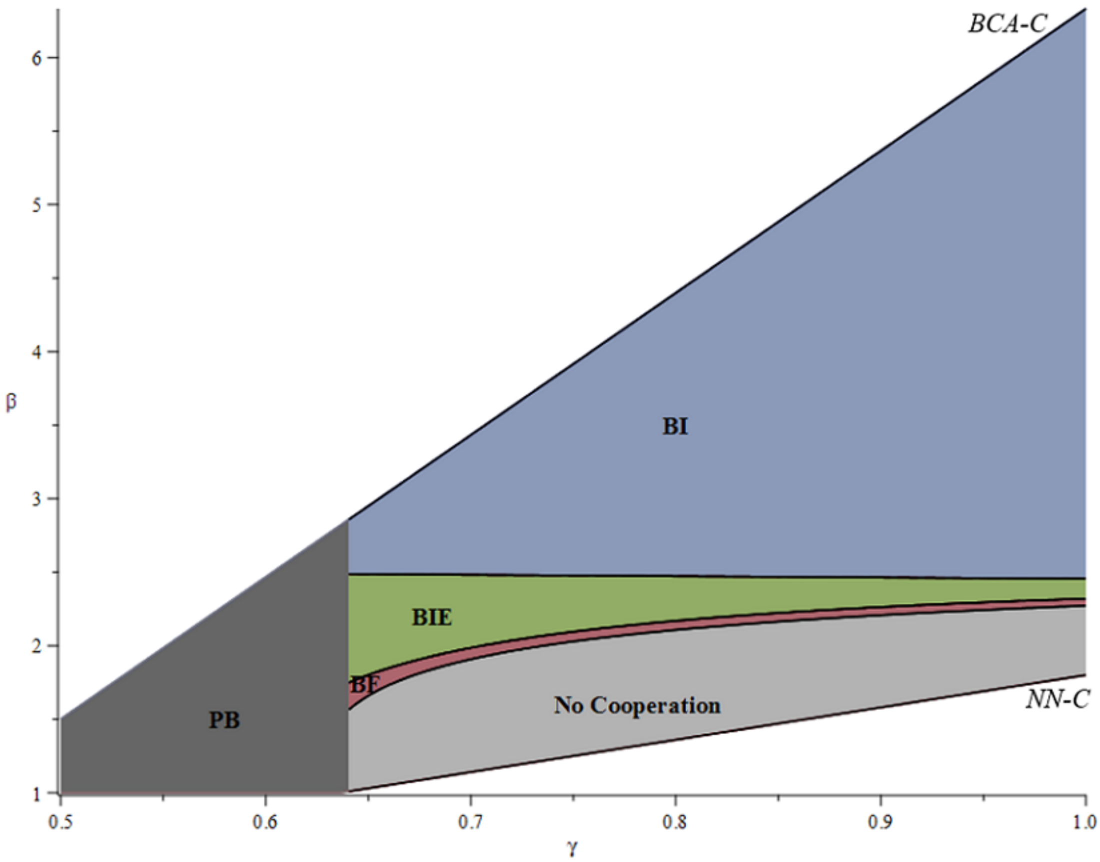


Figure 3: The Effect of BCAs on the Incentive of the Environmentally Less Concerned Country to Cooperate

*BCA-C and NN-C are the BCA-constraint and the non-negativity constraint, respectively.

The second issue that we need to address is the credibility of threats, which takes a look at welfare of country 1. That is, in stage I, country 1 must be better off under full cooperation than under the PB-regime in order for the BI-threat to be credible. With reference to Figure 2, country 1's welfare at node 4 must be higher than at node 3, otherwise country 1 would not threaten with the implementation of the BI-regime. As we know from Proposition 5, country 1 is always better off under full cooperation than under the PB-regime. This piece of information which is repeated for convenience is condition i in Lemma 2 below.

Similarly, suppose the BI-threat is not successful to establish cooperation. Country 1 will use the BIE-threat but only if it is better off under full cooperation than under the BI-regime. That is, country 1's welfare must be higher at node 6 than node 5. This is condition ii in Lemma 2 below.

Moreover, country 1 should be better off under full cooperation than under the BIE-regime for the BF-threat to be credible. Hence, country 1's welfare at node 8 must be higher than at node 7. This is condition iii in Lemma 2 below. Finally, should country 1 not be successful with the BF-threat, it must be better off implementing any of the BCA-threats than under the PB-regime, which we know is true from Proposition 5. Again, this piece of information is repeated for convenience in condition iv in Lemma 2.

Lemma 2. The Credibility of BCA-threats by Country 1

i. Credibility of BI-threat: country 1 is better off under full cooperation than under the PB-regime.

ii. Credibility of BIE-threat: country 1 is better off under full cooperation than under the BI-regime if $\beta \leq \overline{\beta_1}(\gamma)$.

iii. Credibility of BF-threat: country 1 is better off under full cooperation than under the BIE-regime if $\beta \leq \overline{\beta_2}(\gamma)$.

We have: $\overline{\beta_1}(\gamma) > \overline{\beta_2}(\gamma)$.

iv. Country 1 is better off under all three BCA-regimes than under the PB-regime.

Proof. See Appendix B.2 and Proposition 5. □

In contrast to country 2, country 1 is better off under full cooperation than under the BCA-regimes if damages are sufficiently high relative to the net benefits from production and consumption, i.e. if β is not too large. Moreover, Lemma 2 shows that country 1 would have less incentive to fully cooperate if it is equipped with both carbon tariffs and export rebates compared to carbon tariffs only. Note also that the two thresholds $\overline{\beta_1}(\gamma)$ and $\overline{\beta_2}(\gamma)$ under which country 1 prefers full cooperation over the BI- and the BIE-regime increase in γ, i.e. in its individual marginal damages.

Finally, we need to combine Lemma 1 and 2 in order to determine the equilibrium path. For instance, in stage III in Figure 2, we note from Lemma 1 that we only proceeded to stage III because $\beta < \underline{\beta_2}(\gamma)$. That is, previous attempts to enforce cooperation have failed. Suppose, the BF-threat is successful in stage III, i.e., $\beta \geq \underline{\beta_3}(\gamma)$. Hence, we have together $\underline{\beta_2}(\gamma) > \beta \geq \underline{\beta_3}(\gamma)$. From Lemma 2, condition iii, we know that the credibility of the BF-threat requires $\beta \leq \overline{\beta_2}(\gamma)$. Now since we can show that $\underline{\beta_2}(\gamma) < \overline{\beta_2}(\gamma)$, cooperation is an equilibrium path if $\underline{\beta_2}(\gamma) > \beta \geq \underline{\beta_3}(\gamma)$. That is, if the enforcement condition of Lemma 1 for stage III holds, the corresponding credibility condition in stage III holds as well. The same procedure is applied to stage I and II with similar conclusions which are summarised in Proposition 6.

Proposition 6. Subgame-Perfect Nash Equilibrium in the Escalating Penalty Game

1) Stability Region

- *If $\beta \geq \underline{\beta_1}(\gamma)$: full cooperation is achieved along the path: Cooperation→No Cooperation→BI→Cooperation, and the game ends in stage I at node 4 in Figure 2.*

- *If $\underline{\beta_1}(\gamma) > \beta \geq \underline{\beta_2}(\gamma)$: full cooperation is achieved along the path: Cooperation → No Cooperation → BI →No Cooperation →BIE → Cooperation, and the game ends in stage II at node 6 in Figure 2.*

- *If $\underline{\beta_2}(\gamma) > \beta \geq \underline{\beta_3}(\gamma)$: full cooperation is achieved along the path: Cooperation→ No Cooperation → BI →No Cooperation→BIE →No Cooperation→BF→Cooperation, and the game ends in stage III at node 8 in Figure 2.*

2) Instability Region

- *If $\underline{\beta_3}(\gamma) > \beta$, either the BF-regime is implemented following the BF-threat or if country 1 implements its most preferred BCA regime, then either (a) or (b) below:*

  (a) If $\underline{\beta_3}(\gamma) > \beta > 2\gamma$, there is no cooperation and the BIE-regime is implemented with an export rebate which is not a full rebate ($\varphi^* < 1$):

Cooperation→No Cooperation→BI→No Cooperation→BIE, and the game ends in stage III at node 7 in Figure 2.

(b) If $2\gamma \geq \beta$, there is no cooperation and the BI-regime is implemented: Cooperation→No Cooperation→BI, and the game ends in stage II at node 5 in Figure 2.

Proof. See Appendix B.3. □

In the stability region, full cooperation is established based on the threat by country 1 to impose BCA-measures. Hence, there are three paths to reach full cooperation. If $\beta \geq \underline{\beta_1}(\gamma)$, country 2 cooperates as a reaction to the BI-threat. If $\underline{\beta_1}(\gamma) > \beta \geq \underline{\beta_2}(\gamma)$, the BIE-threat works and if $\underline{\beta_2}(\gamma) > \beta \geq \underline{\beta_3}(\gamma)$ only the BF-threat works. Since the potential gains from cooperation decrease in β as we know from Proposition 5, this implies that if the potential gains from cooperation would be large, only harsh punishment works if at all but fails for $\underline{\beta_3}(\gamma) > \beta$.

For $\underline{\beta_3}(\gamma) > \beta$, although country 1 would be better off under full cooperation than any of the BCA-regimes, none of these threats are successful to enforce full cooperation. Hence, country 1 decides which BCA-measure to implement. The instability region is divided into two sub-regions, in which either the BIE-regime or the BI-regime is implemented, depending on the parameter range, or instead, the BF-regime if one argues that the last threat must be implemented. In any case, the three BCA-regimes are preferred to the PB-regime by country 1.

If we considered an alternative penalty game with only one stage in which country 1 moves first and can choose among the three BCA-threats right from the beginning as discussed in Appendix B.4, we find that full cooperation is the subgame-perfect Nash equilibrium following a BF-threat for $\underline{\beta_2}(\gamma) > \beta \geq \beta_3(\gamma)$. However, for $\beta_1(\gamma) > \beta \geq \beta_2(\gamma)$, we have two subgame-perfect equilibria following a BIE- and BF-threat and for $\beta \geq \beta_1(\gamma)$, we have three subgame-perfect equilibria following a BI-, BIE- and BF-threat. Hence, our escalating penalty game in Figure 2 may be considered as an equilibrium refinement to the game in B.4 with a unique equilibrium for each parameter range.

5.3 Normative Analysis of the Cooperation Stage

In this subsection, we briefly evaluate our results from a normative perspective. We showed in Proposition 5 that BCAs, under some conditions, would cause a global welfare loss compared to the non-cooperative PB-regime. Hence, it is important to understand whether the implementation of BCAs threats if they were not successful causes a global welfare loss compared to the PB-regime.

Corollary 1. *Compared to the PB-regime, if BCAs are associated with a global welfare loss, they are used only as threats and are not implemented (stability regions), while if they are implemented they improve global welfare (instability region).*

Proof. Follows from comparing the threshold levels for which BCAs lead to a global welfare loss in Proposition 5 with the threshold levels for which full cooperation is achieved in Proposition 6 where $\underline{\beta}_W^{BF} > \underline{\beta_1} \, \forall \, \gamma$. □

Thus, the negative impact of BCA-measures on global welfare are avoided due to its strategic role to enforce cooperation. However, if they need to be implemented, they lead to higher global welfare. That is, we find that all parameter values for which BCAs lead to a global welfare loss fall in the stability region. However, it is also worthwhile to recall that country 1 has no incentive to implement BCA-measures along any equilibrium path in the stability region because cooperation is preferred to BCA-measures.

Another observation which we made above in Proposition 5 was that the potential gains from cooperation, $\Delta W = W^{FC*} - W^{PB*}$ would be large if β was small. From Proposition 6 we know that if β is sufficiently small, i.e., $\underline{\beta_3}(\gamma) > \beta$, none of the BCA-regimes enforce cooperation. Combining both results, we ask the question to which extent do BCAs when implemented in the instability region close the gap ΔW. In order to answer this question, we employ a relative measure called the closing the gap index (CGI) as suggested by Eyckmans and Finus (2006).

$$CGI_W^{BI} = \frac{W^{BI*} - W^{PB*}}{\Delta W} \; \& \; CGI_W^{BIE} = \frac{W^{BIE*} - W^{PB*}}{\Delta W} \tag{10}$$

Corollary 2. BCAs and the Global Welfare Gap

i. The larger the potential global gains from cooperation, the less effective are BCA-measures to enforce full cooperation.

ii. In the instability region, CGI_W^{BI} and CGI_W^{BIE} are decreasing when lowering β whereas the global welfare gap ΔW is increasing.

Proof. Follows from Proposition 5 and 6 and Appendix B.5. □

On the one hand, BCAs close the global welfare gap fully through their strategic role for all $\beta \geq \underline{\beta_3}$, i.e., in the stability region. However, if β is sufficiently small,

$\beta < \underline{\beta}_3$, full cooperation cannot be achieved. (As the BF-regime leads to lower global welfare than the other two BCA-regimes, it has been ignored in the analysis above.) On the other hand, the lower β, the larger would be the global gains from full cooperation. Hence, whenever cooperation would be needed most, BCAs do not enforce cooperation. Moreover, the larger the potential gains from cooperation are expected to be, the smaller the success of BCAs implementation. This result is in line with Barrett (1994) who referred to this as the 'paradox of cooperation'.[15]

6 Conclusions

The absence of a global policy to mitigate climate change raises concerns about carbon leakage which undermines the effectiveness of unilateral actions. Economists and policy makers have suggested border carbon adjustments (BCAs) as a measure both to address leakage effects and enforce cooperative climate agreements. In this paper, we assessed the effectiveness of three forms of BCAs in an intra-industry trade model and solved a strategic carbon tax competition game between two countries which differ in their perception of environmental damages. Our game comprises three stages: countries play an enforcement game in stage 1, choose their carbon taxes in stage 2 and firms choose their output in stage 3.

We assumed that country 1 is more concerned about environmental damages and gives a higher weight to damages than country 2. In the enforcement stage, we considered the damage evaluation of country 2 not to be too high. In such cases, it becomes worse off under full cooperation with a uniform socially optimal tax than under the non-cooperative PB-regime with bilateral production-based taxes. The benefits of reduced damages from global emission reduction and higher tax revenues does not compensate country 2 for the loss of producer and consumer surplus.

Apart from the PB-regime, we considered three designs of BCAs. a) BCAs on imports, implying that country 1 imposes unilaterally a tariff on imports from country 2, where this tariff fully adjusts the difference between the carbon taxes in the two countries (BI-regime). b) BCAs on imports are complemented by country 1 giving rebates to its firm on its exports to country 2 where the rebate rate is chosen optimally (BIE-regime). c) The same as b, though the export rebate

[15] Barrett (1994) proposed this term in an environmental agreement game without trade. In his context, either only small agreements are stable or if large agreements are stable, then the global gains from cooperation are small.

is not chosen optimally but is a full rebate, which de facto means that country 1 imposes a unilateral consumption-based tax (BF-regime).

We showed that carbon tariffs and export rebates protect the competitiveness of country 1's firm in the home and the foreign market, respectively. Country 1 can better control output and hence global emissions through BCAs as leakage effects are much lower. However, tariffs disadvantage consumers in country 1 and export rebates are costly, which explained that if country 1 can choose its rebate optimally, it may not choose a full rebate. However, possible disadvantages of BCAs for country 1 are offset by the benefits to which also tariff revenues contribute. Essentially, tariffs shift tax revenues from country 2 to country 1.

A comparison across the four non-cooperative regimes, PB-, BI-, BIE- and BF-, showed the following. First, all the BCA-regimes reduce global emissions compared to the PB-regime. Second, the BI-regime with only tariffs reduces global emissions more than the two BCA-regimes with export rebates (BIE- and BF-regime). Third, compared to the social optimum, the BI-regime is the least distortionary regime in global welfare terms should BCAs be implemented. Fourth, BCAs may lead to a global welfare loss compared to the PB-regime, though this occurs only if the global gains from full cooperation are small, i.e., global damages compared to the net benefits from production and consumption are small. The first four results suggest that BCA-measures should be used with caution, should ideally not be implemented and only be used as a threat to enforce cooperation. However, if implemented, the BI-regime is superior to the BIE- and BF-regime. Fifth, country 1 is always better off implementing unilaterally BCA-measures compared to the PB-regime, but for country 2 this is normally not the case. The BF-regime is the least attractive of the three BCA-regimes in terms of global welfare, but also for country 2, which formed the basis for our escalating penalty game in stage 1.

In stage 1, we assumed country 2 will not agree to a fully cooperative agreement as it is worse off than under the non-cooperative PB-regime. We then tested under which conditions the incentives of country 2 will change when faced with the threat by country 1 to implement unilateral BCA-measures. Our results showed that the higher environmental damages are compared to the net benefits from production and consumption, i.e., the larger are the potential gains from cooperation, the harsher must be the threat of punishment by country 1 to enforce cooperation. We showed that the sequence of escalating penalties is deterrent but also credible and proceeds from the BI-, BIE- to the BF-regime. We demonstrated that this sequence is an equilibrium path (subgame-perfect equilibrium) and can be successful in enforcing cooperation. However, whenever the potential gains from full cooperation are expected to be really large, even the harshest punishment fails to establish cooperation. We argued that this has some resemblance with Barrett's paradox of cooperation (Barrett, 1994).

In this paper, we considered one aspect of asymmetry among countries, which was the evaluation of environmental damages. One could also look at other aspects, as for instance different carbon intensities across countries (Fischer and Fox, 2012; Mattoo et al., 2009; Weitzel et al., 2012). Another possible extension could be to extend this model to an n-player asymmetric agreement formation game which tests for coalition stability along the lines in Al Khourdajie and Finus (2018). Finally, one could analyse possible transfer mechanisms between the two heterogeneous countries along the lines of the literature on agreement formation, asymmetric countries and optimal transfers (Finus and McGinty, 2019; Pavlova and De Zeeuw, 2013) by adding trade and BCAs to these models.

References

Al Khourdajie, A. and Finus, M. (2018). Measures to enhance the effectiveness of international climate agreements: the case of border carbon adjustments. Working paper No. 71/18, Bath Economics Research Papers.

Anouliés, L. (2015). The strategic and effective dimensions of the border tax adjustment. *Journal of Public Economic Theory*, 17(6):824–847.

Babiker, M. H. and Rutherford, T. F. (2005). The economic effects of border measures in subglobal climate agreements. *The Energy Journal*, 26(4):99–125.

Baksi, S. and Chaudhuri, A. R. (2017). International trade and environmental cooperation among heterogeneous countries. In Kayalıca, M. Ö., Çağatay, S., and Mıhçı, H., eds., *Economics of International Environmental Agreements*, pp. 97–115. Routledge.

Barnett, A. H. (1980). The Pigouvian tax rule under monopoly. *The American Economic Review*, 70(5):1037–1041.

Barrett, S. (1994). Self-enforcing international environmental agreements. *Oxford Economic Papers*, 46:878–894.

Barrett, S. (1997). The strategy of trade sanctions in international environmental agreements. *Resource and Energy Economics*, 19(4):345–361.

Böhringer, C., Balistreri, E. J., and Rutherford, T. F. (2012a). The role of border carbon adjustment in unilateral climate policy: overview of an energy modeling forum study EMF 29. *Energy Economics*, 34:S97–S110.

Böhringer, C., Carbone, J. C., and Rutherford, T. F. (2012b). Unilateral climate policy design: efficiency and equity implications of alternative instruments to reduce carbon leakage. *Energy Economics*, 34:S208–S217.

Böhringer, C., Carbone, J. C., and Rutherford, T. F. (2016). The strategic value of carbon tariffs. *American Economic Journal: Economic Policy*, 8(1):28–51.

Böhringer, C., Fischer, C., and Rosendahl, K. E. (2014). Cost-effective unilateral climate policy design: size matters. *Journal of Environmental Economics and Management*, 67(3):318–339.

Brander, J. A. and Spencer, B. J. (1985). Export subsidies and international market share rivalry. *Journal of International Economics*, 18(1-2):83–100.

Branger, F. and Quirion, P. (2014). Would border carbon adjustments prevent carbon leakage and heavy industry competitiveness losses? Insights from a meta-analysis of recent economic studies. *Ecological Economics*, 99:29–39.

Copeland, B. R. (1996). Pollution content tariffs, environmental rent shifting, and the control of cross-border pollution. *Journal of International Economics*, 40(3-4):459–476.

Cosbey, A., Droege, S., Fischer, C., and Munnings, C. (2019). Developing guidance for implementing border carbon adjustments: lessons, cautions, and research needs from the literature. *Review of Environmental Economics and Policy*, 13(1):3–22.

Duval, Y. and Hamilton, S. F. (2002). Strategic environmental policy and international trade in asymmetric oligopoly markets. *International Tax and Public Finance*, 9(3):259–271.

Eichberger, J. (1993). *Game Theory for Economists*. Academic Press.

Elliott, J., Foster, I., Kortum, S., Munson, T., Perez Cervantes, F., and Weisbach, D. (2010). Trade and carbon taxes. *American Economic Review*, 100(2):465–69.

Eyckmans, J. and Finus, M. (2006). New roads to international environmental agreements: the case of global warming. *Environmental Economics and Policy Studies*, 7(4):391–414.

Eyland, T. and Zaccour, G. (2012). Strategic effects of a border tax adjustment. *International Game Theory Review*, 14(03):1250016.

Eyland, T. and Zaccour, G. (2014). Carbon tariffs and cooperative outcomes. *Energy Policy*, 65:718–728.

Finus, M. and McGinty, M. (2019). The anti-paradox of cooperation: diversity may pay! *Journal of Economic Behavior and Organization*, 157:541–559.

Fischer, C. and Fox, A. K. (2012). Comparing policies to combat emissions leakage: border carbon adjustments versus rebates. *Journal of Environmental Economics and Management*, 64(2):199–216.

Friedman, J. W. (1986). *Game Theory with Applications to Economics.* Oxford University Press, USA.

Hecht, M. and Peters, W. (2018). Border adjustments supplementing nationally determined carbon pricing. *Environmental and Resource Economics*, 73(1):93–109.

Helm, D., Hepburn, C., and Ruta, G. (2012). Trade, climate change, and the political game theory of border carbon adjustments. *Oxford Review of Economic Policy*, 28(2):368–394.

Hoel, M. (1996). Should a carbon tax be differentiated across sectors? *Journal of Public Economics*, 59(1):17–32.

Irfanoglu, Z. B., Sesmero, J. P., and Golub, A. (2015). Potential of border tax adjustments to deter free riding in international climate agreements. *Environmental Research Letters*, 10(2):024009.

Ismer, R. and Neuhoff, K. (2007). Border tax adjustment: a feasible way to support stringent emission trading. *European Journal of Law and Economics*, 24(2):137–164.

Keen, M. and Kotsogiannis, C. (2014). Coordinating climate and trade policies: Pareto efficiency and the role of border tax adjustments. *Journal of International Economics*, 94(1):119–128.

Kennedy, P. W. (1994). Equilibrium pollution taxes in open economies with imperfect competition. *Journal of Environmental Economics and Management*, 27(1):49–63.

Markusen, J. R. (1975). International externalities and optimal tax structures. *Journal of International Economics*, 5(1):15–29.

Martin, R., Muûls, M., De Preux, L. B., and Wagner, U. J. (2014). Industry compensation under relocation risk: a firm-level analysis of the EU emissions trading scheme. *American Economic Review*, 104(8):2482–2508.

Mattoo, A., Subramanian, A., Van Der Mensbrugghe, D., and He, J. (2009). Reconciling climate change and trade policy. The World Bank.

Pavlova, Y. and De Zeeuw, A. (2013). Asymmetries in international environmental agreements. *Environment and Development Economics*, 18(1):51–68.

Sanctuary, M. (2018). Border carbon adjustments and unilateral incentives to regulate the climate. *Review of International Economics*, 26(4):826–851.

Stiglitz, J. (2006). A new agenda for global warming. *The Economists' Voice*, 3(7).

Tsakiris, N., Michael, M. S., and Hatzipanayotou, P. (2014). Asymmetric tax policy responses in large economies with cross-border pollution. *Environmental and Resource Economics*, 58(4):563–578.

Weitzel, M., Hübler, M., and Peterson, S. (2012). Fair, optimal or detrimental? Environmental vs. strategic use of border carbon adjustment. *Energy Economics*, 34:S198–S207.

Winchester, N. (2011). The impact of border carbon adjustments under alternative producer responses. *American Journal of Agricultural Economics*, 94(2):354–359.

Yonezawa, H., Balistreri, E. J., Kaffine, D. T., et al. (2012). The suboptimal nature of applying Pigouvian rates as border adjustments. Colorado School of Mines: Division of Economics and Business Working Papers of the Center of Economic Research at ETH Zurich.

Appendix

A Appendix of Section 4

In the third stage, the output stage, using (2) and (3), delivers $\frac{\partial^2 \pi_k}{\partial x_{ki}^2} = -2 < \frac{\partial^2 \pi_k}{\partial x_{ki} \partial x_{\ell i}} = -1 < 0$, which guarantees a unique Nash equilibrium in output levels in each market (Eichberger, 1993; Friedman, 1986). That is, $\left(\frac{\partial^2 \pi_k}{\partial x_{ki}^2} \frac{\partial^2 \pi_\ell}{\partial x_{\ell i}^2} \right) - \left(\frac{\partial^2 \pi_k}{\partial x_{ki} \partial x_{\ell i}} \frac{\partial^2 \pi_\ell}{\partial x_{\ell i} \partial x_{ki}} \right) =3 > 0 \; \forall \; k \neq \ell$.

The welfare functions in (7) and (8) can be written explicitly as follows:

$$
\begin{aligned}
W_1 = & \underbrace{\frac{(x_{11}+x_{21})^2}{2}}_{CS_1} + \underbrace{(x_{11})^2 + (x_{12})^2}_{PS_1} + \underbrace{t_1(x_{11}+x_{12})}_{TR_1} - \underbrace{\gamma de}_{D_1} \\
& + \underbrace{(t_1 - t_2)x_{21}}_{BCAI_1} - \underbrace{\varphi t_1 x_{12}}_{BCAE_1}
\end{aligned} \tag{A.1}
$$

$$
W_2 = \underbrace{\frac{(x_{22}+x_{12})^2}{2}}_{CS_2} + \underbrace{(x_{22})^2 + (x_{21})^2}_{PS_2} + \underbrace{t_2(x_{22}+x_{21})}_{TR_2} - \underbrace{(1-\gamma)\, de}_{D_2} \tag{A.2}
$$

We define $\beta = \frac{A}{d}$. We recall, $A = a - c$ is a proxy for the market size or the net benefits of production and consumption and d is the global damage parameter.

Due to lack of space, the subsequent proofs are a sketch; detailed computations are available from the authors upon request.

A.1 FC-Regime

Inserting the uniform taxes $t_{1i} = t_{2i} = t$ for $i = 1, 2$ into (4), gives equilibrium outputs $x_{ki}^{FC} = \frac{A-t}{3}\ \forall i = 1, 2$ and $k = 1, 2$. Inserting equilibrium outputs into the aggregate welfare function $W^{FC} = W_1^{FC} + W_2^{FC}$, differentiating with respect to t yields the following first- and second order condition:

$\frac{\partial W^{FC}}{\partial t} = -\frac{4}{9}(A + 2t) + \frac{4}{3}d = 0$, and $\frac{\partial^2 W}{\partial t^2} = -\frac{8}{9} < 0$. Solving for the socially optimal carbon tax leads to:

$$t^{FC*} = -\frac{1}{2}A + \frac{3}{2}d\,. \tag{A.3}$$

Hence, we have $x_{ki}^{FC*} = (A-d)/2$, $e^{FC*} = 2(A-d)$ and $W_1^{FC*} = \frac{(A-d)(A+d-4\gamma d)}{2}$, $W_2^{FC*} = \frac{(A-d)(A-3d+4\gamma d)}{2}$, and $W^{FC*} = (A-d)^2$. In order to have positive production levels (i.e., an interior solution), we impose the non-negativity constraint (NN- constraint) $A > d$ or, using $\beta = \frac{A}{d}$, $\beta > 1$.

Note that this constraint as well as those required under the other regimes are summarised in Table A.1 at the end of Appendix A.

A.2 PB-Regime

Inserting the effective taxes in Table 1 into (4) leads to: $x_{11}^{PB} = x_{12}^{PB} = \frac{A-2t_1+t_2}{3}$ and $x_{22}^{PB} = x_{21}^{PB} = \frac{A-2t_2+t_1}{3}$. Inserting these outputs into (A.1) and (A.2) and setting $BCAI_1 = BCAE_1 = 0$, give the first-order conditions as follows:

$\frac{\partial W_1^{PB}}{\partial t_1} = -\frac{1}{9}(4A + 7t_1 + t_2) + \frac{2}{3}\gamma d = 0$, and
$\frac{\partial W_2^{PB}}{\partial t_2} = -\frac{1}{9}(4A + 7t_2 + t_1) + \frac{2}{3}(1-\gamma)\,d = 0.$

Solving $\frac{\partial W_1^{PB}}{\partial t_1} = 0$ and $\frac{\partial W_2^{PB}}{\partial t_2} = 0$ simultaneously gives equilibrium taxes:

$$t_1^{PB*} = d\left(\gamma - \frac{1}{8}\right) - \frac{1}{2}A\,, \tag{A.4}$$

$$t_2^{PB*} = d\left(\frac{7}{8} - \gamma\right) - \frac{1}{2}A\,, \tag{A.5}$$

where the second order conditions are satisfied: $\frac{\partial^2 W_i}{\partial {t_i}^2} = -\frac{7}{9} < 0$, $\frac{\partial^2 W_i}{\partial t_i \partial t_j} = -\frac{1}{9} < 0$, and $\frac{\partial^2 W_i}{\partial {t_1}^2}\frac{\partial^2 W_2}{\partial {t_2}^2} - \frac{\partial^2 W_1}{\partial t_1 \partial t_2}\frac{\partial^2 W_2}{\partial t_2 \partial t_1} = \frac{16}{27} > 0$, which ensures a unique Nash equilibrium. These conditions are are also sufficient for the Routh-Hurwitz stability condition to be satisfied (Brander and Spencer, 1985). The slope of the reaction function under this regime is given by $\frac{\partial t_i(t_j)}{\partial t_j} = -\frac{\partial^2 W_i}{\partial t_i \partial t_j} / \frac{\partial^2 W_i}{\partial {t_i}^2} = -\frac{1}{7} < 0$, and reaction functions are given by $RF_1^{PB} := t_1(t_2) = -\frac{1}{7}(4A + t_2 - 6\gamma d)$ and $RF_2^{PB} := t_2(t_1) = -\frac{1}{7}(4A + t_1 - 6(1-\gamma)d)$.

Inserting equilibrium taxes into outputs, we obtain equilibrium outputs, equilibrium welfare levels W_1^{PB*} and W_2^{PB*} with $W^{PB*} = W_1^{PB*} + W_2^{PB*} = \frac{(4A-7d)(4A-d)}{16}$ and $e^{PB*} = 2A - \frac{1}{2}d$. From equilibrium outputs, the NN-constraint under this regime is given by $\beta > \frac{1}{4}(8\gamma - 3)$. Moreover, we always have $t_1^{PB*} > t_2^{PB*}$ for all $\gamma > \frac{1}{2}$.

A.3 BI-Regime

Inserting the effective taxes in Table 1 into (4) leads to: $x_{11}^{BI} = x_{21}^{BI} = \frac{A-t_1}{3}$, $x_{12}^{BI} = \frac{A-2t_1+t_2}{3}$ and $x_{22}^{BI} = \frac{A-2t_2+t_1}{3}$. Inserting these outputs into (A.1) and (A.2) and setting $BCAE_1 = 0$, we obtain W_1^{BI} and W_2^{BI}. The first-order conditions are:

$\frac{\partial W_1^{BI}}{\partial t_1} = -\frac{1}{9}(A + 10t_1 - 2t_2) + \gamma d = 0$, and $\frac{\partial W_2^{BI}}{\partial t_2} = -\frac{1}{3}(t_1 + t_2) + \frac{1}{3}(1-\gamma)d = 0$.

Solving $\frac{\partial W_1^{BI}}{\partial t_1} = 0$ and $\frac{\partial W_2^{BI}}{\partial t_2} = 0$ simultaneously, gives equilibrium taxes:

$$t_1^{BI*} = \left(\frac{7}{12}\gamma + \frac{1}{6}\right)d - \frac{1}{12}A\,, \tag{A.6}$$

$$t_2^{BI*} = \left(\frac{5}{6} - \frac{19}{12}\gamma\right)d + \frac{1}{12}A\,, \tag{A.7}$$

where the second order conditions are: $\frac{\partial^2 W_1}{\partial {t_1}^2} = -\frac{10}{9} < 0$, $\frac{\partial^2 W_1}{\partial t_1 \partial t_2} = \frac{2}{9} > 0$, and $\frac{\partial^2 W_2}{\partial {t_2}^2} = -\frac{1}{3} < 0$, $\frac{\partial^2 W_2}{\partial t_2 \partial t_1} = -\frac{1}{3} < 0$. Therefore, we have $\frac{\partial^2 W_1}{\partial {t_1}^2}\frac{\partial^2 W_2}{\partial {t_2}^2} - \frac{\partial^2 W_1}{\partial t_1 \partial t_2}\frac{\partial^2 W_2}{\partial t_2 \partial t_1} = \frac{4}{9} > 0$, which ensures a unique and stable equilibrium.

The slopes of the reaction functions of countries are given by: $\frac{\partial t_1(t_2)}{\partial t_2} = -\frac{\partial^2 W_1}{\partial t_1 \partial t_2} / \frac{\partial^2 W_1}{\partial {t_1}^2} = \frac{1}{5} > 0$ and $\frac{\partial t_2(t_1)}{\partial t_1} = -\frac{\partial^2 W_2}{\partial t_2 \partial t_1} / \frac{\partial^2 W_2}{\partial {t_2}^2} = -1 < 0$, and the reaction functions by: $RF_1^{BI} := t_1(t_2) = \frac{-A+9\gamma d}{10} + \frac{1}{5}t_2$ and $RF_2^{BI} := t_2(t_1) = (1-\gamma)d - t_1$.

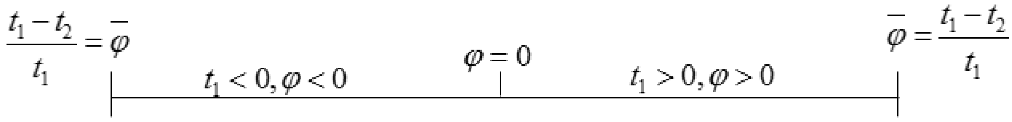
Inserting equilibrium taxes into outputs, it turns out that the most restrictive NN-constraint requires $\beta > \frac{1}{5}(11\gamma - 2)$ and global emissions are given by $e^{BI*} = \frac{25A - d(\gamma - 8)}{18}$.

Since the difference between the two equilibrium taxes is ambiguous, we need to impose a BCA-constraint such that $t_1^{BI*} > t_2^{BI*}$. The BCA-constraint requires $\beta < 13\gamma - 4$. Inserting t_1^{BI*} and t_2^{BI*} into welfare functions, gives W_1^{BI*}, W_2^{BI*}, and $W^{BI*} = W_1^{BI*} + W_2^{BI*}$.

A.4 BIE-Regime

In models with imperfect competition, generally, the equilibrium tax can be positive or negative (in which case it is a subsidy). Therefore, the feasible values of the rebate rate depends on the equilibrium taxes in country 1 and 2. Moreover, we need to consider $t_1 > t_2$ and $t_1(1-\varphi) \geq t_2$.

If $t_1 > 0$, $\varphi > 0$, we have $0 < \varphi \leq \overline{\varphi} = \frac{t_1 - t_2}{t_1}$, where for the maximum allowable rebate rate $\overline{\varphi}$, $\overline{\varphi} \leq 1$ holds if $t_1 > t_2 \geq 0$, while $\overline{\varphi} > 1$ if $t_2 < 0$. If $0 > t_1 > t_2$, then $\varphi < 0$. In such cases, the feasible values for φ is $0 > \varphi \geq \overline{\varphi} = \frac{t_1 - t_2}{t_1}$. This can be illustrated as follows:



Inserting the effective taxes in Table 1 into (4), gives the equilibrium output levels: $x_{11}^{BIE} = x_{21}^{BIE} = \frac{A - t_1}{3}$, $x_{12}^{BIE} = \frac{A - 2t_1(1-\varphi) + t_2}{3}$ and $x_{22}^{BIE} = \frac{A - 2t_2 + t_1(1-\varphi)}{3}$. Inserting these outputs into (A.1) and (A.2) gives the welfare function of each country under this regime.

The first-order conditions are given by:

$\frac{\partial W_1^{BIE}}{\partial t_1} = -\frac{1}{9}\left(A + 10t_1 + 4t_1\varphi^2 - 2t_2 - \varphi\left(A + 8t_1 + t_2 - 3\gamma d\right)\right) + \gamma d = 0\,,$

$\frac{\partial W_1^{BIE}}{\partial \varphi} = \frac{1}{9}t_1\left(A + 4t_1 - 4t_1\varphi + t_2 - 3\gamma d\right) = 0\,,$

$\frac{\partial W_2^{BIE}}{\partial t_2} = -\frac{1}{3}\left(t_1 + t_2\right) + \frac{1}{3}(1-\gamma)d = 0$, which is similar to the previous regime.

Solving $\frac{\partial W_1^{BIE}}{\partial t_1} = 0$, $\frac{\partial W_2^{BIE}}{\partial t_2} = 0$ and $\frac{\partial W_1^{BIE}}{\partial \varphi} = 0$ simultaneously, the Nash equilibrium carbon taxes are given by:

$$t_1^{BIE*} = \frac{1}{3}d\left(\gamma + 1\right) > 0\,, \tag{A.8}$$

$$t_2^{BIE*} = \frac{2}{3}d\left(1 - 2\gamma\right) \leq 0\,, \tag{A.9}$$

and the optimal export rebate rate is given by:

$$\varphi^{*} = \frac{3\left(A + d\left(2 - 3\gamma\right)\right)}{4\,d\left(\gamma + 1\right)} > 0\,, \tag{A.10}$$

noting that $\varphi^{*} \leq (>)1$ if $\beta \leq (>)\ddot{\beta}(\gamma) = \frac{1}{3}\left(13\gamma - 2\right)$, with the second order conditions being satisfied: $\frac{\partial^2 W_1}{\partial {t_1}^2} = -\frac{10}{9} - \frac{4\varphi}{9}(\varphi + 2) < 0\;\forall\,\varphi$, $\frac{\partial^2 W_1}{\partial t_1 \partial t_2} = \frac{2+\varphi}{9} > 0$ for all $\varphi \in (-2, \infty)$, $\frac{\partial^2 W_1}{\partial t_1 \partial t_2} = \frac{2+\theta}{9} < 0$ for all $\varphi \in (-\infty, -2)$ and $\frac{\partial^2 W_2}{\partial {t_2}^2} = -\frac{1}{3} <$

0, $\frac{\partial^2 W_2}{\partial t_2 \partial t_1} = -\frac{1}{3} < 0$, where $\frac{\partial^2 W_1}{\partial t_1{}^2}\frac{\partial^2 W_2}{\partial t_2{}^2} - \frac{\partial^2 W_1}{\partial t_1 \partial t_2}\frac{\partial^2 W_2}{\partial t_2 \partial t_1} = \frac{4}{9} + \frac{\varphi(4\varphi-7)}{27} > 0\, \forall \varphi$ and $\frac{\partial^2 W_1}{\partial \varphi^2} = -\frac{4}{9}t_1^2 < 0\, \forall t_1 \neq 0$.

The slopes of the reaction functions are given by: $\frac{\partial t_1(t_2)}{\partial t_2} = \frac{\varphi+2}{2(2\varphi^2-4\varphi+5)} > 0\, \forall\, \varphi > 0$ and $\frac{\partial t_2(t_1)}{\partial t_1} = -1 < 0$ and the reaction functions by: $RF_1^{BIE} := t_1(t_2) = \frac{t_2(\varphi+2)-A+9\gamma d+\varphi(A-3\gamma d)}{2(2\varphi^2-4\varphi+5)}$, and $RF_2^{BIE} := t_2(t_1) = (1-\gamma)d - t_1$.

Inserting equilibrium taxes into outputs, it turns out that the most restrictive NN-constraint requires $\beta > \frac{1}{3}(7\gamma - 2)$. Global emissions are $e^{BIE*} = \frac{51A - d(5\gamma-14)}{36}$. We also need to impose a BCA-constraint such that $t_1^{BIE*}(1-\varphi^*) \geq t_2^{BIE*}$, which leads to $\beta \leq \frac{1}{3}(29\gamma - 10)$. Note that $t_1^{BIE*} > t_2^{BIE*}$ always holds. Inserting t_1^{BIE*} and t_2^{BIE*} into the two welfare functions, we obtain W_1^{*BIE} , W_2^{*BIE}, and $W^{*BIE} = W_1^{*BIE} + W_2^{*BIE}$.

A.5 BF-Regime

Inserting the effective taxes in Table 1 into (4) gives equilibrium output: $x_{11}^{BF} = x_{21}^{BF} = \frac{A-t_1}{3}$, $x_{12}^{BF} = \frac{A+t_2}{3}$ and $x_{22}^{BF} = \frac{A-2t_2}{3}$. Inserting these outputs into (A.1) and (A.2) gives the welfare function of each country and the first-order conditions are given by:

$\frac{\partial W_1^{BF}}{\partial t_1} = \frac{1}{3}(t_2 - 2t_1 + 2\gamma d) = 0$, and $\frac{\partial W_2^{BF}}{\partial t_2} = -\frac{1}{3}(t_1 + t_2) + \frac{1}{3}(1-\gamma)d = 0$, which is similar to the previous two regimes.

Solving $\frac{\partial W_1^{BF}}{\partial t_1} = 0$ and $\frac{\partial W_2^{BF}}{\partial t_2} = 0$ simultaneously, the equilibrium carbon taxes are given by:

$$t_1^{BF*} = \frac{1}{3}d(\gamma + 1) > 0\,, \tag{A.11}$$

$$t_2^{BF*} = \frac{2}{3}d(1 - 2\gamma) \leq 0\,, \tag{A.12}$$

with the second-order conditions being satisfied: $\frac{\partial^2 W_1}{\partial t_1{}^2} = -\frac{2}{3} < 0$, $\frac{\partial^2 W_1}{\partial t_1 \partial t_2} = \frac{1}{3} > 0$, and $\frac{\partial^2 W_2}{\partial t_2{}^2} = -\frac{1}{3} < 0$, $\frac{\partial^2 W_2}{\partial t_2 \partial t_1} = -\frac{1}{3} < 0$, where $\frac{\partial^2 W_1}{\partial t_1{}^2}\frac{\partial^2 W_2}{\partial t_2{}^2} - \frac{\partial^2 W_1}{\partial t_1 \partial t_2}\frac{\partial^2 W_2}{\partial t_2 \partial t_1} = \frac{1}{3} > 0$. The slopes of the reaction function are given by $\frac{\partial t_1(t_2)}{\partial t_2} = \frac{1}{2} > 0$ and $\frac{\partial t_2(t_1)}{\partial t_1} = -1 < 0$ and the reaction functions by: $RF_1^{BF} := t_1(t_2) = \frac{1}{2}t_2 + \gamma d$ and $RF_2^{BF} := t_2(t_1) = (1-\gamma)d - t_1$.

Inserting equilibrium taxes into outputs, the most restrictive NN-constraint requires $\beta > \frac{1}{3}(\gamma + 1)$. Global emissions are given by $e^{BF*} = \frac{12A + 2d(\gamma-2)}{9}$.

Under this regime, there is no need to impose a BCA-constraint as we always have $t_1^{BF*} > t_2^{BF*}$ and $t_1^{BF*}(1-\varphi) = 0 \geq t_2^{BF*}$ with (=) if only $\gamma = 0.5$.

Inserting equilibrium taxes into welfare functions gives W_1^{BF*} and W_2^{BF*} and $W^{BF*} = W_1^{BF*} + W_2^{BF*}$.

A.6 All Regimes

We summarise the conditions that satisfy the NN-constraint and the BCA-constraint in the following table. For comparisons across regimes, we use the most restrictive condition, which is summarised under "Feasible Range".

Table A.1: Feasible Range of Parameters Values

Regime/Constraint	NN-constraint	BCA-constraint
FC	$\beta > 1$	/
PB	$\beta > \frac{1}{4}(8\gamma - 3)$	/
BI	$\beta > \frac{1}{5}(11\gamma - 2)$	$\beta < 13\gamma - 4$
BIE	$\beta > \frac{1}{3}(7\gamma - 2)$	$\beta \leq \frac{1}{3}(29\gamma - 10)$
BF	$\beta > \frac{1}{3}(\gamma + 1)$	/
Feasible Range	$\beta > \check{\beta} = 1$ for all $\gamma < 0.6363$ $\beta > \check{\beta} = \frac{1}{5}(11\gamma - 2)$ for all $\gamma > 0.6363$	$\beta \leq \hat{\beta} = \frac{1}{3}(29\gamma - 10)$

A.7 Reaction Functions

As shown in Figure 1 (b) and (c), the reaction functions of both countries under the BCA-regimes are piecewise. For instance in Figure 1 panel (b), $RF_1^{BI} := t_1(t_2)$, which is given in Appendix A.3, intersects with the 45°-line at $\bar{t_2} = \frac{1}{8}(9\gamma d - A)$. At this tax level, matching taxes $t_1(\bar{t_2}) = \bar{t_2}$ is a best response of country 1. For any tax level $t_2 \geq \bar{t_2}$, the reaction function jumps to $RF_1^{PB} := t_1(t_2)$, given in Appendix A.2. Similarly for country 2, $RF_2^{BI} := t_2(t_1)$ intersects with the 45°-line at $\bar{t_1} = \frac{1}{2}d(1-\gamma)$. For any tax level $t_1 \leq \bar{t_1}$, country 2's reaction function jumps to $RF_2^{PB} := t_2(t_1)$. Taken together, the constraint needed for $t_2 < \bar{t_2}$ and $t_1 > \bar{t_1}$ is the BCA-constraint stated in Appendix A.3. There are two possibilities for the reaction functions to intersect above the 45°-line: the intersection of the RFs under the PB-regime and the intersection of the RFs under the BI-regime. The first possibility is not possible for all $\gamma \geq 0.5$ which we assume in our model, while the second possibility is ruled out by the BCA-constraint. Hence, the unique solution is the intersection of the RFs under the BI-regime below the 45°-line.

Similar analysis applies to Figure 1 panel (c). As shown above, the RF_2 under the three BCA-regimes is the same, and, hence, we have the same intersection points with the 45°-line as above. For country 1, the reaction function RF_1^{BIE} and RF_1^{BF},which are given in Appendix A.4 and A.5, respectively, intersect with the 45°-line at $\bar{t_2} = \frac{\varphi(A-3\gamma d)+9\gamma d-A}{4\varphi^2-9\varphi+8}$ under the BIE-regime and at $\bar{t_2} = 2\gamma d$ under the BF-regime. For any $t_2 \geq \bar{t_2}$, the reaction function of country 1 jumps to the PB-regime.

By comparing the slopes of the reaction functions of country 1 under the three BCA-regimes, we find that the reaction function of country 1 under the BI-regime is always steeper than under BF-regime and is also steeper than under the BIE-regime for all $\varphi \in \left(0, \frac{13}{4}\right)$, which is satisfied as long as the BCA-constraints hold. The reaction function under the BIE-regime is steeper than under the BF-regime for all $\varphi \in (0, 1)$ and for all $\varphi \in \left(\frac{3}{2}, \infty\right)$, while the reaction function is flatter under the BIE-regime than under the BF-regime for all $\varphi \in \left(1, \frac{3}{2}\right)$. For $\varphi = 1$, they have the same slope.

A.8 Proof of Proposition 3

The ranking of equilibrium tax levels follows directly from comparing the taxes provided in (A.3) to (A.12), using the NN-and BCA-constraints in the feasible range in Table A.1, i.e. $\check{\beta} < \beta \leq \hat{\beta}$. From Appendix A.4, we have $t_1^{BIE*} > 0$ and $\varphi^* > 0$. Thus, the effective tax $t_{12}^{BIE*} = t_1^{BIE*}(1 - \varphi^*) = \frac{d(13\gamma-2)-3A}{12}$. Comparing t_{12}^{BIE*} with t_{12}^{PB*} and t_{12}^{BI*}, which are the equilibrium taxes under the PB-and the BI-regime, we find that $t_{12}^{BI*} > t_{12}^{BIE*}$ and $t_{12}^{BIE*} > t_{12}^{PB*}$ always hold as long as the NN-constraints hold. Whether t_{12} is larger under the BIE- than under the BF-regime depends on the value of the export rebates. Hence, $t_{12}^{BF*} \leq (>)\, t_{12}^{BIE*}$ if $\varphi^* \leq (>) 1$, where t_{12}^{BF*} is de facto zero. Finally, we could have $t_{12}^{BF*} \leq t_{12}^{PB*}$ if $\beta \leq 2\gamma - \frac{1}{4}d$. However, this condition violates the constraint $\beta > \check{\beta} = 1$ for all $\gamma \leq 0.625$ and violates the constraint $\beta > \check{\beta} = \frac{1}{5}\left(11\gamma - 2\right)$ for all $\gamma \geq 0.75$. Hence, if $\beta \leq 2\gamma - \frac{1}{4}d$ for all $0.625 < \gamma < 0.75$, we have $t_{12}^{BF*} < t_{12}^{PB*} < t_{12}^{BIE*} < t_{12}^{BI*}$ because in such cases, $\varphi^* < 1$.

A.9 Proof of Proposition 4

We compare global emission levels which are given in Appendix A.2 to A.5. Recall that we use below the NN- and BCA-constraints in the "Feasible Range" listed in Table A.1 above.

- Comparison with the PB-regime:

  $e^{PB*} < e^{BI*}$ if $\beta < \frac{1}{11}\left(1 - \gamma\right)$, $e^{PB*} < e^{BIE*}$ if $\beta < \frac{1}{21}\left(4 - 5\gamma\right)$, and $e^{PB*} < e^{BF*}$ if $\beta < \frac{1}{3}\left(\gamma + \frac{1}{4}\right)$, where all the above conditions can be shown to violate the NN-constraints. Hence, we have $e^{BI*}, e^{BIE*}, e^{BF*} < e^{PB*}$.

- Across the BCA-regimes:

  $e^{BIE*} \leq e^{BI*}$ if $\beta \leq 3\gamma - 2$, which violates the NN-constraints. Hence, we have $e^{BIE*} > e^{BI*}$.

$e^{BF*} \leq e^{BI*}$ if $\beta \geq 5\gamma$, which violates the BCA-constraint only if $\gamma < 0.714$. Note that if $\beta \geq 5\gamma$, this implies that $\varphi^* > 1$.

$e^{BIE*} \leq (>) e^{BF*}$ if $\beta \leq (>) \ddot{\beta}(\gamma) = \frac{1}{3}(13\gamma - 2)$, i.e., if the optimal rebate is less than or equal (larger than) a full rebate, i.e., $\varphi^* \leq (>) 1$.

A.10 Proof of Proposition 5

Using Appendices from A.1 to A.5, and upon substitution of equilibrium taxes and outputs in country's welfare function, equilibrium welfare is obtained.

1) The global welfare gap between the FC- and the PB-regime is $\Delta W = W^{FC*} - W^{PB*} = \frac{9}{16}d^2$.

2) BCA-regimes vs PB-regime

- $W^{BI*} \leq W^{PB*}$ if $\beta \geq$ (or $\leq$) $\frac{1}{61}(\gamma + 152) + (-)\frac{9}{122}\sqrt{1034 - 256\gamma(\gamma - 1)}$. The first condition violates the BCA-constraint for all $\gamma \leq 0.853$, while the second condition violates the NN-constraints. Therefore, $W^{BI*} > W^{PB*}$ except if $\beta > \underline{\beta}_W(BI) = \frac{1}{61}(\gamma + 152) + \sqrt{1034 - 256\gamma(\gamma - 1)}$ for all $\gamma > \gamma_1 = 0.853$.

- $W^{BIE*} \leq W^{PB*}$ if $\beta \geq$ (or $\leq$) $\frac{1}{75}(202 - 23\gamma) + (-)\frac{3}{25}\sqrt{434 - 16\gamma(2 + \gamma)}$. The first condition violates the BCA-constraint for all $\gamma \leq 0.843$, while the second condition is not feasible. Therefore, $W^{BIE*} > W^{PB*}$ except if $\beta > \underline{\beta}_W(BIE) = \frac{1}{75}(202 - 23\gamma) + \frac{3}{25}\sqrt{434 - 16\gamma(2 + \gamma)}$ for all $\gamma > \gamma_2 = 0.843$.

- $W^{BF*} \leq W^{PB*}$ if $\beta \geq$ (or $\leq$) $\frac{1}{3}(\gamma + 7) + (-)\frac{1}{4}\sqrt{81 - 16\gamma^2}$. The first condition violates the BCA-constraint for all $\gamma \leq 0.83$, while the second condition violates the NN-constraints. Therefore, $W^{BF*} > W^{PB*}$ except if $\beta > \underline{\beta}_W(BF) = \frac{1}{3}(\gamma + 7) + \frac{1}{4}\sqrt{81 - 16\gamma^2}$ for all $\gamma > \gamma_3 = 0.83$.

  We also have $\underline{\beta}_W(BI) > \underline{\beta}_W(BIE) > \underline{\beta}_W(BF)$ for all $\gamma > 0.58$, and from the results above, a global welfare loss under any of the BCA-regimes can only be incurred if $\gamma > 0.83$.

- For country 1, $W_1^{BI*} \leq W_1^{PB*}$ if $\beta \leq$(or $\geq$)$\frac{5}{13}(17\gamma - 2) - (+)\frac{3}{26}\sqrt{64\gamma(49\gamma - 10) - 113}$ for all $\gamma > 0.317$. The first inequality violates the NN-constraints and the second inequality violates the BCA-constraint. Therefore, we have $W_1^{BI*} > W_1^{PB*}$. Similar results are obtained by comparing the welfare level of country 1 under the PB-regime with the BIE- and BF-regime. That is, $W_1^{BIE*} > W_1^{PB*}$ if $\frac{(25\gamma-2)}{3} - \Psi < \beta < \frac{(25\gamma-2)}{3} + \Psi$ with

$\Psi = \frac{\sqrt{3}\sqrt{848\gamma^2-128\gamma-31}}{6}$, and $W_1^{BF*} > W_1^{PB*}$ if $\frac{2(11\gamma-1)}{3} - \Omega < \beta < \frac{2(11\gamma-1)}{3} + \Omega$ with $\Omega = \frac{\sqrt{800\gamma^2-128\gamma-31}}{4}$, where these conditions must hold given our NN- and BCA-constraints.

- For country 2, $W_2^{BI*} > W_2^{PB*}$ if and only if $\psi - \xi < \beta < \psi + \xi$, where $\psi = \frac{1}{11}(91 - 127\gamma)$ and $\xi = \frac{9}{44}\sqrt{2}\sqrt{32\gamma(49\gamma - 71) + 843}$. The NN-constraints guarantee the satisfaction of the first part of the above condition $\psi - \xi < \beta$, which does not violate the BCA-constraint. With respect to the second part of the above condition, we have the following: *a)* the BCA-constraint assures the satisfaction of this condition $\beta < \psi + \xi$ for all $\gamma \in [0.5, 0.6275]$, *b)* for $\gamma \in (0.6275, 0.72503)$ country 2 might be better off if and only if the second part of the above condition holds, i.e. if $\beta < \psi + \xi$, *c)* for all $\gamma \in [0.72503, 1]$, this condition violates the NN-constraint. Hence, $W_2^{BI*} > W_2^{PB*}$ for all $\gamma \in [0.5, 0.6275]$, while $W_2^{BI*} < W_2^{PB*}$ for all $\gamma \in [0.72503, 1]$. For the other two BCA-regimes, we have similar conditions, for instance $W_2^{BIE*} > W_2^{PB*}$ for all $\gamma \in [0.5, 0.59]$, while $W_2^{BIE*} < W_2^{PB*}$ for all $\gamma \in [0.7, 1]$, and $W_2^{BF*} > W_2^{PB*}$ for all $\gamma \in [0.5, 0.595]$, while $W_2^{BF*} < W_2^{PB*}$ for all $\gamma \in [0.655, 1]$.

3) Across BCAs regimes:

- $W^{BIE*} \geq W^{BI*}$ if $\beta \leq 3\gamma - 2$ or if $\beta \geq \frac{1}{19}(89\gamma + 42)$. The first condition violates the NN-constraints, while the second condition violates the BCA-constraint. Therefore, $W^{BIE*} < W^{BI*}$. In addition, we find that $W^{BF*} < W^{BI*}$ for all values of γ.

- $W^{BF*} > W^{BIE*}$ if $\frac{1}{3}(13\gamma - 2) = \ddot{\beta} < \beta < \frac{1}{21}(19\gamma + 58)$ for all $\gamma < 1$. The first part of the inequality implies that the optimal rebate is larger than a full rebate $\varphi^* > 1$, and the second inequality is satisfied by the BCA-constraint for all $\gamma < 0.7$. Therefore, $W^{BF*} \leq W^{BIE*}$ if $\beta \leq \ddot{\beta}$, i.e. if $\varphi^* \leq 1$, and if $\beta \geq \frac{1}{21}(19\gamma + 58)$ for all $\gamma \geq 0.7$.

B Appendix of Section 5

B.1 Proof of Lemma 1

i. From Appendix A.1 and A.3: $W_2^{FC*} \geq W_2^{BI*}$ if $\beta \geq \underline{\beta_1}(\gamma) = \frac{1}{11}(28 - \gamma)$ and/or if $\beta \leq 14 - 23\gamma \ \forall \ \gamma \geq 0.5$. The NN-constraints are not sufficient to guarantee the satisfaction of the first condition, thus it needs to hold. However, this condition violates the BCA-constraint for all $\gamma < 0.6025$. Recall that we consider only the range in which cooperation cannot be achieved under the PB-regime, i.e., $\gamma > \overline{\gamma} = 0.6406$ from Proposition 5. Therefore, the first condition does not violate the BCA-constraint in our range. The second condition violates the NN-constraint for all $\gamma \geq 0.57$. Hence, this condition is not relevant for the parameter range which we consider $\gamma > \overline{\gamma}$. As a result, if country 1 imposes the BI-threat, we have $W_2^{FC*} \geq W_2^{BI*}$ if $\beta \geq \underline{\beta_1}(\gamma)$, where $\frac{\partial \underline{\beta_1}}{\partial \gamma} < 0$ for all γ.

ii. From Appendix A.1 and A.4: $W_2^{FC*} \geq W_2^{BIE*}$ if (a) $\gamma < 0.59$ and (b) if $\beta \geq$ (or $\leq$)$\frac{226-323\gamma}{39} + (-)\frac{4\sqrt{2}\sqrt{353\gamma^2-263\gamma+32}}{13}$ for all $\gamma \geq 0.59$. We consider the range: $\gamma > \overline{\gamma} = 0.6406$. The NN-constraints are not sufficient to guarantee the first condition in (b) and it also does not violate the BCA-constraint. As a result, this condition needs to hold. The second condition in (b) violates the NN-constraints. Therefore, for the range $\gamma > \overline{\gamma}$, if country 1 imposes the BIE-threat, $W_2^{FC*} \geq W_2^{BIE*}$ if $\beta \geq \underline{\beta_2}(\gamma) = \frac{226-323\gamma}{39} + \frac{4\sqrt{2}\sqrt{353\gamma^2-263\gamma+32}}{13}$, where $\frac{\partial \beta_2}{\partial \gamma} > 0$ for all $\gamma > 0.59$.

iii. $W_2^{FC*} \geq W_2^{BF*}$ if (a) $\gamma < 0.627$ and (b) if $\beta \geq$ (or $\leq$)$\frac{1}{3}(16 - 20\gamma) + (-)\sqrt{\gamma(46\gamma - 40) + 7}$ for all $\gamma \geq 0.627$. We consider the range: $\gamma > \overline{\gamma} = 0.6406$. The NN-constraints are not sufficient to guarantee the first condition in (b) and it also does not violate the BCA-constraint, thus this condition needs to hold. The second condition in (b) violates the NN-constraints for all $\gamma \geq 0.628$ and hence is not relevant here. Therefore, for the range $\gamma > \overline{\gamma}$, if country 1 imposes the BF-threat, $W_2^{FC*} \geq W_2^{BF*}$ if $\beta \geq \underline{\beta_3}(\gamma) = \frac{1}{3}(16 - 20\gamma) + \sqrt{\gamma(46\gamma - 40) + 7}$, where $\frac{\partial \beta_3}{\partial \gamma} > 0$ for all $\gamma > 0.627$.

In addition, we have $\underline{\beta_1}(\gamma) > \underline{\beta_2}(\gamma) > \underline{\beta_2}(\gamma)$ for all $\gamma > \overline{\gamma}$.

B.2 Proof of Lemma 2

i. As mentioned in Proposition 5, $W_1^{FC*} \geq W_1^{PB*}$ for all $\gamma \geq \frac{23}{64} \simeq 0.36$, which must hold as we assume $\gamma \geq 0.5$.

ii. From Appendix A.1 and A.3: $W_1^{FC*} \geq W_1^{BI*}$ if (a) $\beta \leq \overline{\beta_1}(\gamma) = \frac{1}{13}(\gamma + 32)$ and/or (b) if $\beta \geq 13\gamma - 4$ for all $\gamma \geq 0.5$. The inequality in (a) does not violate the NN-constraints, but the BCA-constraint is not sufficient to guarantee this condition for all $\gamma \geq 0.6$. Hence, for all $\gamma > \overline{\gamma}$, the inequality in (a) needs to hold. The inequality in (b) violates the BCA-constraint. Therefore, $W_1^{FC*} \geq W_1^{BI*}$ if $\beta \leq \overline{\beta_1}(\gamma)$, where $\frac{\partial \overline{\beta_1}}{\partial \gamma} > 0$.

iii. From Appendix A.1 and A.4: $W_1^{FC*} \geq W_1^{BIE*}$ if $\beta \leq$ (or $\geq$)$\frac{1}{3}(25\gamma - 2) - (+)\frac{2}{3}\sqrt{3}\sqrt{53\gamma^2 - 44\gamma + 11}$. The first condition does not violate the NN-constraints and can be satisfied as long as the BCA-constraint holds if $\gamma \leq 0.548$. However, for all $\gamma > \overline{\gamma}$, the BCA-constraint is not sufficient and this condition needs to hold. The second condition violates the BCA-constraint for all γ. Therefore, for all $\gamma > \overline{\gamma}$, we have $W_1^{FC*} \geq W_1^{BIE*}$ if $\beta \leq \overline{\beta_2}(\gamma) = \frac{1}{3}(25\gamma - 2) - \frac{2}{3}\sqrt{3}\sqrt{53\gamma^2 - 44\gamma + 11}$, where $\frac{\partial \overline{\beta_2}}{\partial \gamma} > 0$.

iv. See Appendix A.10 point 2.

B.3 Proof of Proposition 6

Using Lemma 1 and Lemma 2, we solve the game by backward induction. We start with the stability region, i.e., cooperation can be established with one of the BCA-threats and then consider the instability region, i.e., cooperation cannot be established. Recall that the escalating penalty game starts from $\gamma > \overline{\gamma}$ (see Assumption 1).

1) Stability Region

In stage III, country 2 faces the BF-threat and can either cooperate or not. Country 2 cooperates if $\beta \geq \underline{\beta_3}$ and ends at node 8 in Figure 2. Country 1 will only use the BF-threat to establish cooperation if it is better off under cooperation than if it implemented the BIE-regime earlier (on node 7). That is, we must have: $W_1^{FC*} \geq W_1^{BIE*}$, which is true if $\beta \leq \overline{\beta_2}(\gamma)$. Finally, country 1 would only implement the BIE-regime if the BIE-threat did not lead to cooperation. That is, $\beta < \underline{\beta_2}$. In other words, the game has progressed to stage III. Thus, in stage III, in order for cooperation to be an equilibrium path, we need $\underline{\beta_2} > \beta \geq \underline{\beta_3}$ and $\beta \leq \overline{\beta_2}(\gamma)$. Since we have $\underline{\beta_2} < \overline{\beta_2}$, cooperation is an equilibrium path if $\underline{\beta_2} > \beta \geq \underline{\beta_3}$.

In stage II, country 2 faces the BIE-threat and can either cooperate or not. Country 2 chooses no cooperation if $\beta < \underline{\beta_2}$, and we end up in Stage III as described above. Instead, country 2 chooses cooperation in stage II if $\beta \geq \underline{\beta_2}$, and we end up in node 6. Country 1 will only use the BIE-threat to establish cooperation if it is better off under cooperation than if it implemented the BI-regime earlier on (node 5). That is, we must have: $W_1^{FC*} \geq W_1^{BI*}$, which is true if $\beta \leq \overline{\beta_1}(\gamma)$. Finally, country 1 would only implement the BI-regime if the BI-threat did not lead to cooperation. That is, $\beta < \underline{\beta_1}$. This means the game has progressed to stage II. Thus, in stage II, in order for cooperation to be an equilibrium path, we need $\underline{\beta_1} > \beta \geq \underline{\beta_2}$ and $\beta \leq \overline{\beta_1}(\gamma)$. Since we have $\underline{\beta_1} < \overline{\beta_1}$, cooperation is an equilibrium path if $\underline{\beta_1} > \beta \geq \underline{\beta_2}$.

In stage I, country 2 faces the BI-threat and can either cooperate or not. Country 2 chooses no cooperation if $\beta < \underline{\beta_1}$, and we end up in stage II as described above. Instead, country 2 chooses cooperation in stage I if $\beta \geq \underline{\beta_1}$, and we end up in node 4 in Figure 2. Country 1 will only use the BI-threat to establish cooperation if it is better off under cooperation than if it implemented the PB-regime earlier on (node 3). That is, we must have: $W_1^{FC*} \geq W_1^{PB*}$, which we know it always holds (see Lemma 2). Finally, country 1 would only implement the BI-threat if country 2 did not accept its proposal for cooperation in stage 0, which was our starting point as we assume $\gamma > \overline{\gamma}$. Thus, in stage I, in order for cooperation to be an equilibrium path, we need $\beta \geq \underline{\beta_1}(\gamma)$ and $\gamma > \overline{\gamma}$.

2) Instability Region

First, we have shown in Proposition 5 and Lemma 2 that country 1 is better off under any of the BCA-regimes than under the PB-regime. Hence, node 3 is never an equilibrium outcome in the instability region.

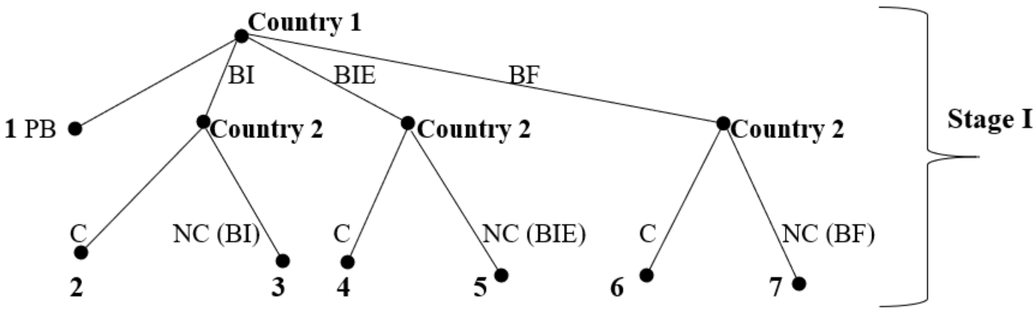
Second, solving $W_1^{BIE*} \geq W_1^{BI*}$ gives two conditions: (a) $\beta \leq 3\gamma - 2$ and (b) $\beta \geq 2\gamma$. The first condition violates the NN-constraints. The second condition does not violate the BCA-constraint, but the NN-constraints are not sufficient to guarantee this condition, thus it needs to hold. Therefore, $W_1^{BIE*} > (\leq) W_1^{BI*}$ if $\beta > (\leq) 2\gamma$.

Third, the BF-regime is dominated by the BIE-regime, and, hence, if $\underline{\beta_3} > \beta > 2\gamma$, country 1 chooses the BIE-regime, while if $\beta \leq 2\gamma$, country 1 chooses the BI-regime.

Note that if $\beta = 2\gamma$, country 1 is indifferent between the BI- and the BIE-regime. Therefore, we assume that country 1 chooses the BI-regime according to the Pareto-criterion. That is, we find $W_2^{BIE*} > W_2^{BI*}$ if and only if $\frac{7}{29}(\gamma - 6) < \beta < (3\gamma - 2)$. The first inequality is satisfied as long as $\beta > 0$, however, the second inequality violates the NN-constraints. Therefore, $W_2^{BIE*} < W_2^{BI*}$.

B.4 One-stage Penalty Game

Suppose stages I, II, and III do not take place sequentially according to the escalating penalty path depicted in Figure 2, but are reduced to stage I as in the Figure below. That is, country 1 can choose either of the three BCA-regimes to enforce cooperation. That is, the game tree would be as follows:



As mentioned in the text, country 1 starts to use the BCA-threats if country 2 refuses the proposal 'cooperation' in stage 0 which is our starting point as in Figure 2. That is, $W_1^{FC*} > W_1^{PB*}$, $W_1^{BI*} > W_1^{PB*}$, $W_1^{BIE*} > W_1^{PB*}$ and $W_1^{BF*} > W_1^{PB*}$, as shown in Appendix A.10 and B.2, and $W_2^{FC*} < W_2^{PB*}$ if $\gamma > \overline{\gamma}$.

If $\beta \geq \underline{\beta_1}(\gamma)$, cooperation (C) is established by the three BCA-measures and we have three subgame-perfect Nash equilibria: BI→C, BIE→C and BF→C, corresponding to endnodes 2, 4 and 6. If $\underline{\beta_1}(\gamma) > \beta \geq \underline{\beta_2}(\gamma)$, we have two subgame-perfect equilibria: BIE→C and BF→C, corresponding to endnodes 4 and 6 and finally for $\underline{\beta_2} > \beta \geq \underline{\beta_3}(\gamma)$, the subgame-perfect equilibrium is BF→C with endnode 6. For $\beta < \underline{\beta_3}$, the outcome is no cooperation (NC), with endnode 5 if $\underline{\beta_3} > \beta > 2\gamma$ because $W_1^{BIE*} > W_1^{BI*}$, whereas if $2\gamma \geq \beta$, endnode 3 would emerge because $W_1^{BIE*} \leq W_1^{BI*}$. Thus, only for $\beta < \underline{\beta_2}$ would we have a unique equilibrium, but not for $\beta \geq \underline{\beta_2}$. However, qualitatively, results would be the same as in Proposition 6.

B.5 Proof of Corollary 2

i. As shown in Appendix A.10, the global welfare gap, $\Delta W = \frac{9}{16}d^2$, increases in d at an increasing rate. We showed in Proposition 6 that full cooperation cannot be achieved if $\underline{\beta_3}(\gamma) > \beta$, i.e., if $\frac{A}{d}$ is low or equivalently if d is high given any value of A.

ii. Inserting the global welfare levels from Appendix A into (10), we obtain $CGI_W^{BIE} = \frac{d^2\left(268\gamma - 73\gamma^2 - 226\right) - 138Ad\left(\gamma - \frac{202}{23}\right) - 225A^2}{1458d^2}$,
where $\frac{\partial CGI_W^{BIE}}{\partial d} < 0$ if $\gamma < \frac{202d-75A}{23d}$ and $\frac{202d-75A}{23d} > 1$ if $\beta < \frac{179}{75} \simeq 2.386$. The BIE-regime is implemented if $\beta < \underline{\beta_3}$, and we have $\underline{\beta_3} < \frac{179}{75}$ for all γ. Hence, $\frac{\partial CGI_W^{BI}}{\partial d} < 0$ in the instability region.

Similarly, we have $CGI_W^{BI} = \frac{d^2((160\gamma - 170\gamma^2 - 71) + 4Ad(\gamma + 152) - 122A^2}{729d^2}$, where $\frac{\partial CGI_W^{BI}}{\partial d} < 0$ if $\gamma > \frac{61A-152d}{d}$, and $\frac{61A-152d}{d} < 0$ if $\beta < \frac{152}{61} \simeq 2.492$. The BI-regime is implemented if $\beta \leq 2\gamma$, where $2\gamma < \frac{152}{61}$ for all γ. Thus, we always have $\frac{\partial CGI_W^{BI}}{\partial d} < 0$ in the instability region.

Part V

Essay 4: Strategic Climate Policies with Endogenous Plant Location: The Role of Border Carbon Adjustments

Appendix 6B: Statement of Authorship

This declaration concerns the article entitled:			
Strategic Climate Policies with Endogenous Plant Location: The Role of Border Carbon Adjustments			
Publication status (tick one) **Draft manuscript** ☒ **Submitted** ☐ **In review** ☐ **Accepted** ☐ **Published** ☐			
Publication details (reference)			
Copyright status (tick the appropriate statement) I hold the copyright for this material ☒ Copyright is retained by the publisher, but I have been given permission to replicate the material here ☐			
Candidate's contribution to the paper (provide details, and also indicate as a percentage)	The candidate contributed to / considerably contributed to / predominantly executed the… Formulation of ideas: - Predominantly contributed to the formulation of ideas. (70%) Design of methodology: - Predominantly contributed to the design of methodology. (85%) ~~Experimental work:~~ Presentation of data in journal format: - Predominantly contributed to the presentation of data in journal format. (65%)		
Statement from Candidate	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature.		
Signed	Noha Nagi Elboghdadly	**Date**	4/10/2019

Last update: Feb 2019

Strategic Climate Policies with Endogenous Plant Location: The Role of Border Carbon Adjustments

Noha Elboghdadly* and Michael Finus†

Abstract

Carbon leakage and the relocation of firms is one of the main concerns of governments when choosing ambitious climate policy measures. In a strategic trade model with endogenous plant location, we study the effect of border carbon adjustments (BCAs) on equilibrium emission taxes in a non-cooperative policy game between two asymmetric countries. For this, we compare a No-BCA with a BCA regime for two scenarios: a simultaneous and a sequential game associated with the notions of Nash equilibrium and Stackelberg equilibrium, respectively. Without BCAs, a 'race to the bottom' is the Nash equilibrium. In a Stackelberg equilibrium, a second less negative 'chicken equilibrium' may emerge, which constitutes a Pareto-improvement not only for the leader but also the follower. In this chicken equilibrium, the Stackelberg leader gives in, letting his/her firms relocate to face a higher tax than in the race-to-the-bottom equilibrium. With BCAs, the race-to-the-bottom in carbon taxes can be avoided in the Nash equilibrium, even though global emissions are higher and global welfare is lower than in the social optimum. However, a Nash equilibrium under the BCA regime may not exist due to the discontinuity and multivaluedness of best response functions, even though it exists if the potential gains from cooperation would be large. In contrast, Stackelberg equilibria always exist, but due to the strategic interaction among countries, BCAs may not always constitute a Pareto-improvement, though they always reduce global emissions and in most cases increase global welfare.

Keywords: Endogenous Plant Location, Global Emissions, Emission Tax Competition, Border Carbon Adjustments

JEL-Classification: R3, F18, F12, Q58, H23, H87

*Department of Economics, University of Bath, 3 East, Bath, BA2 7AY, UK. Email: n.m.w.elboghdadly@bath.ac.uk

†Department of Economics, University of Graz, Universitätsstraße 15, 8010 Graz, Austria and University of Bath, 3 East, Bath, BA2 7AY, UK. Email: michael.finus@uni-graz.at

1 Introduction

The history of climate change negotiations suggests that it is difficult to implement effective measures to reduce greenhouse gases significantly. Effective global actions are hampered by free-rider incentives and sub-global actions are not effective as they are undermined by 'carbon leakage'. That is, emission reductions by some environmentally concerned countries are partly or completely offset by higher emissions in environmentally less concerned countries. One important channel of carbon leakage is the relocation of the production of emission-intensive industries to countries with laxer environmental policies. This phenomenon has been known as the 'pollution haven hypothesis'(PHH). Although there is mixed empirical evidence to support the PHH, the threat of relocation of firms, associated with the loss of jobs and investment, is an important argument in the policy debate.[1]

Recently, border carbon adjustments (BCAs) have been proposed to address the concern of carbon leakage in general and the concern of a loss of competitiveness of domestic industries in particular (Böhringer et al., 2012; Branger and Quirion, 2014; Fischer and Fox, 2012; Stiglitz, 2006; Wooders et al., 2009). Typically, BCAs comprise an import tariff or an export rebate or both. Even ignoring strategic considerations by adopting a pure welfare perspective, trade measures, complementing environmental policies, can be justified as already demonstrated by Markusen (1975). He shows that in the absence of global action, the internalisation of the externality caused by a global pollutant requires a combination of a Pigouvian tax and import tariffs.[2] That is, BCAs correct distortions and hence are not considered as disguised trade barriers (Helm et al., 2012).

In this paper, we are interested in whether and under which conditions BCAs can support the implementation of more ambitious climate policies. Our model takes into account that firms cannot only relocate parts of their production but even their entire production facilities abroad (endogenous plant location) and that governments engage in a strategic emission tax competition game (bilateral and endogenous policy choices) and may perceive global damages from greenhouse gases differently (asymmetric countries).

We model an emission tax competition game between two governments that strive

[1]For example, Eskeland and Harrison (2003) and Manderson and Kneller (2012), among others, find no evidence of the PHH. In contrast, Fredriksson et al. (2003), Xing and Kolstad (2002) and Kellenberg (2009) report significant effects of environmental policies on the location of firms.

[2]See also Hoel (1996) and Copeland (1996) for similar results. These models assume perfect competition.

to attract the plants of two firms in an intra-industry trade model. Both firms produce a homogeneous emission-intensive good and compete in a Nash-Cournot fashion. Governments evaluate damages from global emissions differently. The game comprises three stages: in stage 1, governments choose their policies; in stage 2, firms choose their location and in stage 3 firms choose their output. We solve our game under two different policy regimes. Under the No-BCA regime, each government imposes a carbon tax on the production of those plants, which are located within its national boundaries. Under the BCA regime, the country that sets a higher carbon tax can additionally impose a tariff on imports from plants located abroad. Export rebates are not considered. As mentioned by Markusen et al. (1995), endogenous plant location may lead to a discontinuity of welfare functions with respect to tax levels as firms may change their location abruptly above a threshold. In our model, this leads to non-continuous best response functions under the BCA regime, which implies that a Nash equilibrium may not exist. Therefore, as Stackelberg equilibria always exist, we also consider the possibility that governments choose their taxes sequentially. Apart from this technical point, a sequential policy choice also gives rise to new interesting results.

Under the No-BCA regime, we find that if countries choose their policies simultaneously, we end up in a race-to-the-bottom with no relocation of firms. Different from papers that assume local pollution, this is the only pure strategy Nash equilibrium in our model, irrespective of the degree of asymmetry among countries. In contrast, if governments move sequentially, the Stackelberg leader may be able to avoid being stuck at the bottom. The leader may act as a 'wise chicken', imposing a higher carbon tax than the follower so that all firms move to the follower (total relocation of firms). This Stackelberg equilibrium is Pareto-improving not only for the leader but also the follower and leads to lower global emissions than in the Nash equilibrium.

Under the BCA regime, we show that a Nash equilibrium may not exist, though it exists when the potential gains from cooperation are expected to be large. If the Nash equilibrium exists, a BCA-policy is an effective measure to eliminate the race-to-the-bottom, resulting in higher global welfare and lower global emissions, though global welfare falls short and global emissions exceed those in the social optimum. The location equilibrium entails partial relocation of plants from the more to the less environmentally concerned country. Both countries adjust their taxes upward compared to the race-to-the-bottom equilibrium. The more concerned country imposes a higher emission tax, which is complemented by BCAs such as to avoid the total relocation of its firm.

The effect of BCAs on Stackelberg equilibria depends greatly on the identity of the leader and the parameters values of our model. However, one common result is that a race-to-the-bottom becomes less likely to emerge as an equilibrium. Furthermore, global emissions are reduced, global welfare usually increases, even though the country on which BCAs are imposed may not always be better off.

Our paper is related to two strands of the literature on strategic environmental-trade policies, which all build on the strategic imperfect-competition trade model due to Brander and Spencer (1985) by adding environmental damages (and sometimes consumer utility) in governments welfare function.

The first strand of the literature studies strategic environmental policies assuming fixed plant location (immobile firms). This includes for instance Conrad (1993), Barrett (1994) and Kennedy (1994). This literature concludes for Cournot-competition that if environmental policy is the only instrument available to governments, environmental taxes are set below marginal damages. That is, governments have an incentive to provide their firms with a strategic advantage over their rivals. In Eyland and Zaccour (2014), Anouliés (2015), Baksi and Chaudhuri (2017), and Hecht and Peters (2018), BCAs are added to the tax competition game. In a two-country model, Eyland and Zaccour (2014) show, based on numerical simulations, that BCAs allow countries to set higher carbon taxes in the non-cooperative equilibrium. This has been confirmed by Hecht and Peters (2018) for the country which imposes BCAs, while the country on which BCAs are imposed responds by a lower carbon tax. Anouliés (2015) and Baksi and Chaudhuri (2017) demonstrate that BCAs are helpful as a threat to enforce cooperation among asymmetric countries. The environmentally less concerned country is more likely to engage in a cooperative agreement as its non-cooperative outside payoff is reduced through BCAs.

The second strand of the literature assumes mobile firms, and analyses the effect of environmental policies on the location of firms. Two types of models have emerged: the market share game and the location game where the difference lies in the sequence of the game.[3] In market share games, firms choose first their location and then governments choose their policies, also called ex-post policy game. By construction, governments cannot affect the location of their firms and hence these type of games are less interesting for our analysis.[4] In contrast, in location games, governments move first and then firms choose their location, also

[3]See for instance Ulph and Valentini (2001) and Petrakis and Xepapadeas (2003) for a detailed comparison between the two types of games.

[4]For examples of market share type of games, see for instance Eerola (2006), De Santis and Stähler (2009) and Dijkstra et al. (2011).

called ex-ante policy game. Our model belongs to this class of models.

Some of the early studies of location games include for instance Markusen et al. (1993) and Motta and Thisse (1994). However, these papers, similar to the more recent paper by Sanna-Randaccio et al. (2017), assume an exogenous policy level, and, hence, ignore the effect of firm mobility on the incentives of governments to set their environmental policies strategically. A simple extension to address this shortcoming are plant location games that assume a unilateral endogenous climate policy, as for instance in Petrakis and Xepapadeas (2003) and Ikefuji et al. (2016). However, governments hardly choose their policies in isolation. Hence, a more sophisticated extension allows for endogenous bilateral policy choices similar to that in Markusen et al. (1995), Rauscher (1995), Hoel (1997), and Ulph and Valentini (2001). For instance, Markusen et al. (1995) shows that there are two possible Nash equilibria, which depend on the evaluation of the damages from local pollution: i) a low evaluation leads to low environmental regulation with a 'race to the bottom', and ii) a high evaluation leads to strict environmental regulation with a 'race to the top', also called ' not in my backyard', implying that firms exit the market or moving to other regions. Both Rauscher (1995) and Hoel (1997) simplify the assumptions made by Markusen et al. (1995) by ignoring transportation and set-up costs, and obtain similar qualitative results but do not have to rely on simulations.

All of the papers based on location games do not consider BCAs, most assume local pollution[5], and a monopolistic market structure. Clearly, none of these assumptions is useful in our context as BCAs are proposed to achieve two objectives. First, they aim at internalising a global externality (environmental objective). Second, they aim at reducing leakage effects by levelling the playing field for domestic and foreign firms (competitiveness objective). Therefore, we assume a global pollutant, implying that a not-in-my-backyard argument is not rational for governments. In fact, the argument may just be reversed: keeping firms in the own backyard may be a rational policy if firms face lower environmental taxes abroad. Moreover, in order to capture the competitiveness argument, we assume an oligopolistic market structure.

The remainder of the paper is organised as follows. In Section 2, we present our model, discuss important features and solve the second and third stage of the model. In Sections 3 and 4, we solve the first stage of our game and derive the climate policy equilibria under the two alternative policy regimes, respectively. In Section 5, we compare equilibria in order to show the effect of BCAs. Section 6

[5] Rauscher (1995) considers the possibility of transboundary pollution as an extension.

concludes and discusses possible future research.

2 Model

First, we present the model. Subsequently, we discuss the main features and assumptions of the model. Then we briefly comment on possible location equilibria. Finally, we derive the socially optimal tax levels as a normative benchmark, against which we compare non-cooperative equilibria in our subsequent analysis.

2.1 Basic Ingredients

We consider two countries, respectively, two governments, $i = 1, 2$, which interact strategically. The game unfolds in three stages, which is solved by backward induction. In the first stage, governments choose their policy levels, in the second stage firms choose their location and in the third stage firms choose their outputs.

In the last stage, there are two firms, $k = 1, 2$, which produce a homogeneous good x and compete in outputs in a Cournot-fashion. Firm 1 is initially located in country 1, and firm 2 is initially located in country 2. Each firm has two plants, one plant supplying the home market in country i, the other plant supplying the foreign market j. Markets are segmented. The inverse demand function in market i is given by:

$$p_i(X_i) = a - X_i \quad \forall\ i = 1, 2\,, \tag{1}$$

where p_i is the market price in market i and parameter $a > 0$ is the chock-off price. $X_i = x_{1i} + x_{2i}$ is total consumption in country i where x_{1i} and x_{2i} are the outputs supplied by firm 1 and 2 to market i, respectively. For simplicity, we assume identical firms with a linear production cost function, i.e., $C_{ki}(x_{ki}) = cx_{ki}$ for $k = 1, 2$ and $i = 1, 2$.

Good x is an emission intensive good such as cement or steel which generates greenhouse gas emissions, for example, due to its use and combustion of energy in the production process. Without loss of generality, we assume a constant emission-output ratio across firms, which we normalise to 1, such that an emission tax is de facto an output tax. Hence, profits of firm 1 and 2 obtained in market i are given by

$$\pi_{1i} = (p_i(X_i) - c - t_{1i})x_{1i} \ \&\ \pi_{2i} = (p_i(X_i) - c - t_{2i})x_{2i} \ \forall\ i = 1, 2\,, \tag{2}$$

where t_{1i} is the effective tax which firm 1 faces on its supply to market i and t_{2i} is the effective tax which firm 2 faces on its supply to market i.

We consider two policy regimes: the No-BCA regime and the BCA regime. In order to illustrate the difference between the two regimes, suppose firm k produces with one of its plants for market i. This plant faces an effective tax t_{ki}. Now there are two possible location choices. 1) The plant locates in country i, and, hence, faces tax $t_{ki}= t_i$. 2) The plant locates in country j . Under the No-BCA regime, the plant will simply face the tax imposed by country j, i.e., $t_{ki}= t_j$. Under the BCA regime, the same is true as long as $t_i \leq t_j$. However, if $t_i > t_j$, then under the BCA regime, this firm faces the effective tax $t_{ki} = t_j + \omega(t_i - t_j)$ on its exports to country i, with ω the border tax adjustment parameter (Eyland and Zaccour, 2014).

The simultaneous maximisation of profits obtained in market i by both firms (π_{1i} is maximised with respect to output x_{1i} and π_{2i} with respect to x_{2i}), gives equilibrium quantities (denoted by an asterisk) supplied by firm 1 and 2 in market i

$$x^*_{1i} = \frac{A - 2t_{1i} + t_{2i}}{3} \;\&\; x^*_{2i} = \frac{A - 2t_{2i} + t_{1i}}{3}\,,\ \forall\ i = 1, 2\,, \tag{3}$$

with $A = a - c$, which we interpret as a market size parameter. It is also a proxy for the net benefits from production and consumption. Clearly, output levels need to be non-negative. We will test this later for each location equilibrium. Each firm takes separate quantity decisions for the supply to the home and foreign market. Accordingly, profits obtained in market i are given by:

$$\pi^*_{1i} = (x^*_{1i})^2 \;\&\; \pi^*_{2i} = (x^*_{2i})^2\ \forall\ i = 1, 2\,. \tag{4}$$

Thus, the final stage of the three stage game is a Nash equilibrium in output levels in each of the two markets. As both firms are assumed to be identical in all respects, and, as will be explained below, there are neither fixed nor transportation costs, different profits only stem from differences in effective taxes that firms face.

In the second stage, firms choose their location for each of their two plants simultaneously. That is, they take a decision for each market separately. Generally speaking, firm k supplying market i, compares its profit from locating in country i $\pi_{ki}(i)$ with its profit locating in country j $\pi_{ki}(j)$. As this comparison will generally depend on where the competitor firm ℓ locates, the comparison for market i is based on computing $\Delta\pi_{ki} = \pi_{ki}(i, \ell) - \pi_{ki}(j, \ell)$, $\ell = i, j$, with the first entry in brackets indicating the location of firm k, and the second entry the location of the competitor firm ℓ. For a given location of firm ℓ , firm k will locate in country

i if $\Delta\pi_{ki} > 0$ and will locate in country j if $\Delta\pi_{ki} < 0$. In case of indifference, $\Delta\pi_{ki} = 0$, we assume that a firm´s plant locates in the country of origin. The equilibrium location choice implies mutual best replies by firm k and ℓ with respect to their plants supplying market i. That is, the solution of the second stage is a Nash equilibrium of location choices of plants supplying a particular market. As each firm has two plants, each firm takes two location decisions.

In the first stage, governments choose the level of their emission/output tax t_i based on the following welfare function:

$$W_i = CS_i + PS_i + T_i - D_i + BCA_i\,, \tag{5}$$

where CS_i is the consumer surplus in country i, with the consumer surplus being given by $CS_i = \frac{X_i^2}{2}$ which follows from (1), recalling that the total supply to market i is given by $X_i = x_{1i} + x_{2i}$. PS_i is the producer surplus, which is equal to the sum of profits of plants located in country i. T_i is the tax revenue of government i where $T_i = t_i X^i$ and X^i is the sum of output levels produced in country i. D_i are damages from pollution that are released in the production of good x. We assume a global pollutant and hence damages in country i depend on total emissions, E, regardless of the location of the source of emissions. Hence, damages in country i are $D_i(E)$, $E = X = \sum X^i = \sum X_i$. That is, as we normalise the emission-output coefficient to 1, global emissions are equal to total production, which is equal to total consumption. More specifically, we assume:

$$D\left(E\right) = dE, \quad D_1 = \gamma D(E), \quad D_2 = (1-\gamma)D(E), \gamma \in [0.5, 1]\ , \tag{6}$$

with $d > 0$ a damage parameter, reflecting global marginal damages. Hence, country 1 suffers a portion γ and country 2 a portion $(1-\gamma)$ of global damages where we allow for the possibility that countries perceive or evaluate those damages differently. Given $\gamma \in [0.5, 1]$, country 1 is at least as concerned as country 2 about environmental damages and normally more whenever γ is strictly larger than 0.5. This gives us two benchmarks: a) γ= 0.5 implies a symmetric damage evaluation in both countries and b) γ= 1 implies that country 2 is not concerned at all about environmental damages.

Finally, the last term in the welfare function, BCA_i, stands for the tariff revenues obtained from a border carbon adjustment policy. This term is different for our two policy regimes. Under the No-BCA regime, this term is zero by assumption. Under the BCA regime, this term is positive for the government which imposes a tariff but zero for the other government. Generally,

$BCA_i = \omega(t_i - t_j)[x_{ki}(j) + x_{\ell i}(j)]$ if and only if $t_i > t_j$, otherwise $BCA_i = 0$. That is, generally, one plant or two plants supplying market i could be located in country j.

We assume $\omega = 1$, not only for simplicity but for two other reasons. First, any value of ω above 1 would not be compatible with the equal treatment rule under the World Trade Organization (WTO). Second, any value smaller than 1 would not be optimal for country i if it has the option to use BCAs.[6]

2.2 Basic Features of the Model

In this subsection, we discuss the main assumptions of the model with a closer look at the welfare components mentioned above.

First, we assume that consumption takes place in the two countries that strategically interact, and, hence, consumers matter in our model. This is not only important because a crucial feature of BCAs, which are import tariffs, is their negative impact on consumers, but, even more fundamentally, without consumers, there are no imports on which BCAs could be applied.[7]

Second, BCAs are mainly proposed to internalise global externalities. Hence, we assume a global and not a local pollutant.

Third, we assume that profits of plants located in a country as well as the tax revenues obtained from these plants matter for governments. This is in line with Ulph and Valentini (2001) and Petrakis and Xepapadeas (2003), who assume that profits and tax revenues go to the country in which production takes place. A wider interpretation is that profits are an indicator of the importance of domestic production and associated jobs. This captures the main argument put forward by governments and lobby groups in favour of not implementing a too ambitious climate policy.[8] Moreover, excluding profits from governments´ welfare functions implies that tax competition between governments would only be driven by tax

[6] It is straightforward to show that if country i was to choose ω endogenously, it would choose a value strictly larger than 1. See also Hecht and Peters (2018).

[7] Under the No-BCA regime and if production was sold to a third market, the incentive to set lax environmental standards/low taxes would be reduced in our model. See, for example, Ulph and Valentini (2001).

[8] Some papers do not consider profits in the welfare function of governments or assume full repatriation of profits. For instance, Markusen et al. (1995) assume that profits are distributed throughout the world, whereas Rauscher (1995) considers that profits accrue to a foreign investor. Both Motta and Thisse (1994) and Eerola (2006) assume that profits accrue to the country in which the headquarter of a company is located. Clearly, excluding profits from governments' welfare functions would weaken the incentives to set low taxes. See Janeba (1998) and Ulph and Valentini (2001) on this point.

revenue considerations of governments, which is probably not very plausible, as mentioned by Rauscher (1995). Furthermore, the race-to-the-bottom phenomenon could not be explained for other types of environmental regulations such as environmental standards which do not generate revenues to governments.

Fourth, we assume that location choices of firms depend only on tax differentials. That is, we abstract from transportation and set-up costs. This is not only because of their obvious effects,[9] but also because they do not affect the qualitative results, though they would add greatly to the complexity of the analysis.[10]

Fifth, revenues from border carbon adjustments, i.e., the term BCA_i in welfare function (5), only appears under the BCA regime and only in the welfare function of the country that imposes a higher carbon tax. Given our assumption that BCAs fully adjust the difference between the two taxes (i.e., the effective tax of firm k being located in country j and supplying market i, $t_{ki} = t_j + \omega(t_i - t_j)$, is simply $t_{ki} = t_i$ for $\omega = 1$), they can make a difference to the home market i. Not only the home firm´s but also the foreign firm´s supply to the home market i faces the same tax t_i provided $t_i > t_j$. This implies that all plants supplying country i are subject to the same carbon tax, irrespective of the location of production. In other words, BCAs partially protect the home firm´s profit by levelling the playing field in market i. It is only partial protection because in the foreign market j, the foreign firm will have a competitive advantage over the home firm provided $t_i > t_j$. Of course, the home firm can circumvent this disadvantage by relocating its plant to the foreign country j for the supply of this market (and hence facing t_j instead of t_i for its supply to market j).

Finally, as it will become evident, modeling plant location as an endogenous choice of firms poses some analytical difficulties. It causes not only discontinuous location choices of firms as a function of taxes, but, more importantly, may also causes best response functions of governments to be discontinuous, which, in our model, may lead to the non-existence of a Nash equilibrium under the BCA regime. It is for this reason that apart from Nash equilibria, we also determine Stackelberg equilibria in this first stage of our three stage game. Moreover, for expositional simplicity, we consider only the possibility of a unilateral BCA-policy that is

[9] Fixed costs or set-up costs of plants reduce the incentive of plant relocation. In contrast, transportation costs increase the incentive of relocation of the plant that supplies the foreign market. For a detailed analysis of the effect of these costs on plant location, see for instance Markusen et al.(1993;1995), Motta and Thisse (1994) and Sanna-Randaccio et al. (2017).

[10] For instance, Markusen et al. (1995) consider transportation and set-up costs and hence rely on numerical simulations. In contrast, Rauscher (1995) and Hoel (1997) abstract from those costs and obtain similar analytical results.

imposed by country 1, the country which is more concerned about environmental damages in our model. The general possibility of a bilateral BCA-policy where country 2 could also impose a tariff if $t_i < t_j$ is considered in Appendix E where it is shown that all qualitative conclusions continue to hold.

2.3 Location Equilibria

Firms choose the location of their plants in the second stage based on the carbon taxes chosen by governments in the first stage. Given that we abstract from transportation and set-up costs, the decision of firms depends only on tax differentials as demonstrated in more details in Appendix A.

Under the No-BCA regime, only the location of production matters, which gives rise to three location equilibria. 1) 'No relocation' (NR) if $t_1 = t_2$. Each firm remains with its two plants in the country of origin. 2) Total relocation of firm 1 (TR_1) if $t_1 > t_2$. Firm 1, originally located in country 1, will relocate with both plants to country 2. 3) Total relocation of firm 2 (TR_2) if $t_1 < t_2$. Firm 2, originally located in country 2, will relocate with both plants to country 1.

Under the BCA regime, assuming that only government 1 can impose BCAs, location equilibria NR if $t_1 = t_2$ and TR_2 if $t_1 < t_2$ are the same. However, if $t_1 > t_2$, the TR_1-location equilibrium disappears in favour of the PR_1-location equilibrium, standing for 'partial relocation of firm 1'. Firm 1´s plant supplying market 2 will relocate to country 2, but its plant supplying its own market in country 1 will remain in the country of origin, as also the foreign firm 2 faces de facto the same tax t_1 on its exports to the market in country 1 (and, as pointed out above, we assume that in case of indifference plants do not relocate). That is, country 1 imposing BCAs on imports can avoid total relocation of its firm 1, but cannot avoid partial relocation. If we considered that also country 2 can impose BCAs on imports if $t_1 < t_2$, as we do in Appendix E, then also location equilibrium PR_2 would exist.

2.4 Normative Benchmark

Before turning to non-cooperative equilibria under the two policy regimes, we briefly discuss the normative benchmark of the social optimum. Maximising $W_1 + W_2$ with respect to output levels delivers X_1^{S*} and X_2^{S*}, the socially optimal output levels supplied to market 1 and 2, with $X_1^{S*} = x_{11}^{S*} + x_{21}^{S*}$ and $X_2^{S*} = x_{12}^{S*} + x_{22}^{S*}$. The composition of X_1^{S*} and X_2^{S*} does not matter, as we assume linear and identical production costs for all plants. Moreover, due to a global pollutant (and hence only aggregate damages matter in the social optimum) and because of symmetric consumers, $X_1^{S*} = X_2^{S*}$ must be true.

In order to determine the socially optimal tax, we proceed in two steps. First, the social optimum cannot be achieved with BCAs. Suppose BCAs are imposed by country i with $t_i > t_j$, implying PR_k-location equilibrium. Using equilibrium output levels (3), gives $X_i^* = 2(A - t_i)/3$ for market i and $X_j^* = 2(A - t_j)/3$ for market j from which it is evident that $X_i^{S*} = X_j^{S*}$ is impossible. Second, without BCAs, $X_1^{S*} = X_2^{S*}$ is possible under three location equilibria: a) NR with $t_1 = t_2 = t_S^*$. b) TR_1 with $t_1 > t_2 = t_S^*$ and c) TR_2 with $t_S^* = t_1 < t_2$. All three location equilibria imply de facto the same effective tax rate t_S^* imposed on all firms.

Proposition 1. *Social Optimum*

The socially optimal output levels are given by $X_1^{S*} = X_2^{S*} = A - d$. Under the No-BCA regime, the socially optimal effective tax is given by

$$t_S^* = -\frac{1}{2}A + \frac{3}{2}d \ . \tag{7}$$

with associate output levels $x_{11}^{S*} = x_{12}^{S*} = x_{21}^{S*} = x_{22}^{S*} = (A - d)/2$. A BCA regime cannot generate socially optimal output levels and hence global welfare will be strictly lower.

Hence, the socially optimal effective tax rate is unique with a unique output vector, though it is associated with three different possible location equilibria and tax vectors. In the following, it is helpful to think of the socially optimal tax as a uniform tax imposed in both countries for simplicity. It is interesting to note that - as we will demonstrate later - BCAs can increase global welfare in a non-cooperative equilibrium. However, as Proposition 1 states, BCAs will always fall short of achieving the socially optimal global welfare level. Finally, note that because output levels need to be non-negative, we assume henceforth $A > d$ to ensure interior solutions. From equilibrium output levels in (3), it is evident that a necessary condition to ensure positive production levels requires $A > t_i$, $i \in \{1, 2\}$. We will use these conditions in the subsequent analysis.

3 Climate Policy Equilibria: No-BCA Regime

In this section, we solve the first stage of our game under the No-BCA regime. Based on subsection 2.3, there are three possible location equilibria: 1) NR if $t_1 = t_2$, i.e., all firms remain in their country of origin; 2) TR_1 if $t_1 > t_2$, i.e., all plants are located in country 2; 3) TR_2 if $t_1 < t_2$, i.e., all plants are located in country 1. We consider two equilibria: Nash equilibrium (NE) and Stackelberg equilibrium (SE). In order to determine equilibria, we first derive the best response function of each country (Subsection 3.1) and then use those functions to predict equilibria if countries move simultaneously (Subsection 3.2) and sequentially (Subsection 3.3). Finally, we compare non-cooperative equilibria and contrast them with the social optimum.

3.1 Best Responses

We proceed in three steps. First, we write down the welfare function of each country under the three possible location equilibria. Second, we analyse the best response of each country for a given location equilibrium. Third, we derive the best response of each country across all possible location equilibria.

Based on the general welfare function (5), the welfare function in the three location equilibria can be written as follows:

$$W_i = \begin{cases} W_i^{TR_\ell} = CS_i + \sum\limits_{k=1,2} \pi_{k1} + \sum\limits_{k=1,2} \pi_{k2} + T_i - D_i & \text{if } t_i < t_j \quad (8a) \\ W_i^{NR} = CS_i + \pi_{k1} + \pi_{k2} + T_i - D_i & \text{if } t_i = t_j \quad (8b) \\ W_i^{TR_k} = CS_i - D_i & \text{if } t_i > t_j \quad (8c) \end{cases}$$

where the superscript TR_ℓ refers to the total relocation of the foreign firm ℓ from country j to country i, TR_k refers to the total relocation of the home firm k to country j and NR denotes no relocation. It is evident that in the TR_ℓ-location equilibrium, profits of all four plants accrue to country i, in the NR-location equilibrium, country i enjoys the profits of its domestic firm k with its two plants and in the TR_k-location equilibrium, no profits accrue to country i. Also, the other components (e.g., CS_i and D_i) may be different across the three location equilibria due to different taxes. This is in particular evident for tax revenues, T_i , which in (8a) are given by $T_i^{TR_\ell} = t_i \left(\sum\limits_{k=1,2} x_{k1} + \sum\limits_{k=1,2} x_{k2} \right)$ and in (8b) by $T_i^{NR} = t_i (x_{k1} + x_{k2})$ and are zero in (8c). Inserting equilibrium output levels as provided in Appendix B.1 into (8a), (8b), and (8c) gives (9),(10) and (11), respectively:

$$W_i^{TR_\ell} = \frac{2}{3}\left(A - t_i\right)\left(A + t_i - 2D_i^{'}\right) \; , \tag{9}$$

$$W_i^{NR} = \frac{4}{9}\left(A - t_i\right)\left(A + \frac{1}{2}t_i - 3D_i^{'}\right) \; , \tag{10}$$

$$W_i^{TR_k} = \frac{1}{2}\left(\frac{2}{3}A - \frac{2}{3}t_j\right)^2 - D_i^{'}\left(\frac{4}{3}\left(A - t_j\right)\right) \; , \tag{11}$$

with $D_i^{'}$ the individual marginal damage in country i, i.e., $D_1^{'} = \gamma d$ and $D_2^{'} = \left(1-\gamma\right)d$.

In the TR_ℓ-location equilibrium, the optimal tax of country i follows from differentiation of (9) with respect to t_i, which gives:

$$\frac{\partial W_i^{TR_\ell}}{\partial t_i} = \frac{4}{3}\left(D_i^{'} - t_i\right) = 0 \Leftrightarrow \hat{t}_i = D_i^{'} : \hat{t}_1^{TR_2} = \gamma d \text{ and } \hat{t}_2^{TR_1} = \left(1-\gamma\right)d \tag{12}$$

noting that the second order conditions are satisfied. Given $t_i < t_j$, country i has both firms, can control all output levels, and its optimal unconstrained carbon tax is equal to its marginal damage. On the one hand, country i has an incentive to subsidise its consumers to correct for market distortions due to imperfect competition (Barnett, 1980). On the other hand, and, different from the standard profit-shifting argument, the government has an incentive to tax producers. Country i has both firms and hence no rents can be shifted. Moreover, profits are maximised by taxing producers to enforce the monopolistic output. These two incentives cancel out in our model. The remaining incentive is to internalise environmental damages, which calls for a tax equal to marginal damages in country i.

We note that if $t_j > D_i^{'}$, the unconstrained equilibrium tax of country i is $\hat{t}_i = D_i^{'}$ and the corresponding welfare level of country i is $\hat{W}_i^{TR_\ell}$. However, we also need to consider the possibility that $t_j \leq D_i^{'}$. In this case, country i cannot choose its unconstrained optimum $\hat{t}_i$ but must deviate to $\tilde{t}_i = t_j - \varepsilon$, with ε being positive but close to zero, and the corresponding welfare level is $\tilde{W}_i^{TR_\ell}$. The ε-undercutting constrained optimum follows from the strict concavity of welfare function (9) with respect to t_i.

In the NR-location equilibrium, we just note that $t_i = t_j = t$, which means that welfare can be expressed alternatively as a function of level t_i or t_j. Moreover, W_i^{NR} is strictly concave in t.

Finally, in the TR_k-location equilibrium, we note that the consumer surplus as well as damages are decreasing in t_j and $W_i^{TR_k}$ is a strictly convex function in t_j. Summarising, we need to consider three cases for each country.

Welfare Levels of Country** i **for Three Location Equilibria:

1. Case $t_i < t_j$

   (a) Let $\hat{W}_i^{TR_\ell}$ be $W_i^{TR_\ell}\left(t_i = \hat{t}_i^{TR_\ell}\right)$ in the range $t_j > \hat{t}_i^{TR_\ell} = D_i'$. Country i chooses its unconstrained optimal tax and has both firms.

   (b) Let $\tilde{W}_i^{TR_\ell}$ be $W_i^{TR_\ell}\left(t_i = \tilde{t}_i = t_j - \varepsilon\right)$ as a function of t_j in the range $t_j \leq \hat{t}_i^{TR_\ell} = D_i'$. Country i chooses its constrained optimal tax by marginally undercutting the tax of country j in order to have both firms, with $\varepsilon > 0$ and ε being arbitrarily small and close to zero.

2. Case $t_i = t_j$: let W_i^{NR} be $W_i^{NR}(t_i = t_j)$ as a function of t_j and each firm remains in its home country.

3. Case $t_i > t_j$: let $W_i^{TR_k}$ be $W_i^{TR_k}(t_i > t_j)$ as a function of t_j and both firms are located in country j.

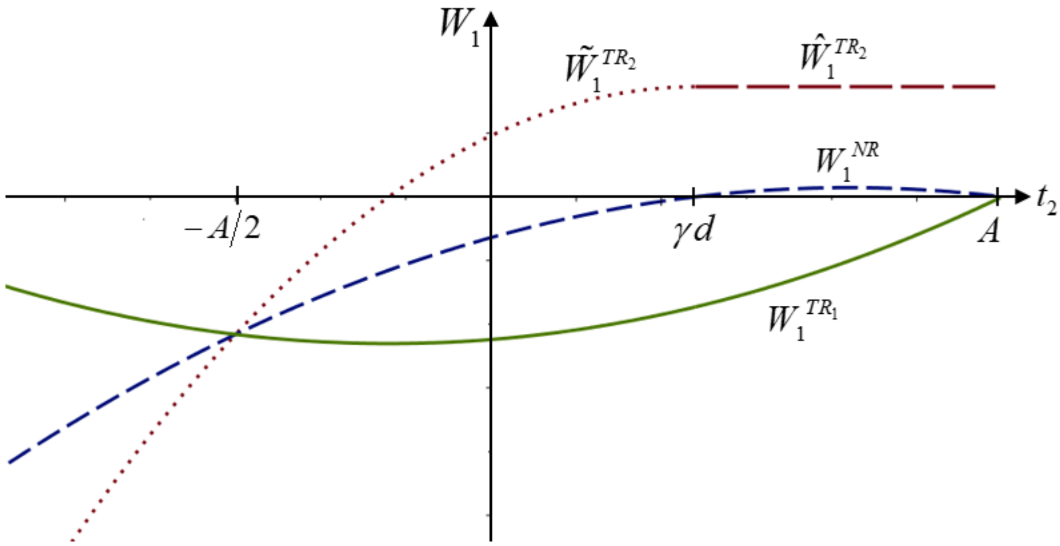
Using the tax levels listed in Case 1 to 3 above, together with (9), (10) and (11), we can compare the welfare of country i for different location equilibria and different levels of t_j, which is illustrated in Figure 1. (The mathematical details are provided in Appendix B.1.)

In Figure 1, part (a) captures the incentives of country 1, and part (b) of country 2. We notice that the upper bound of the parameter space is $t_2 = A$ in part (a) and $t_1 = A$ in part (b) due to the necessary condition for positive production levels. Since part (a) and (b) are very similar, we focus on part (a) first.

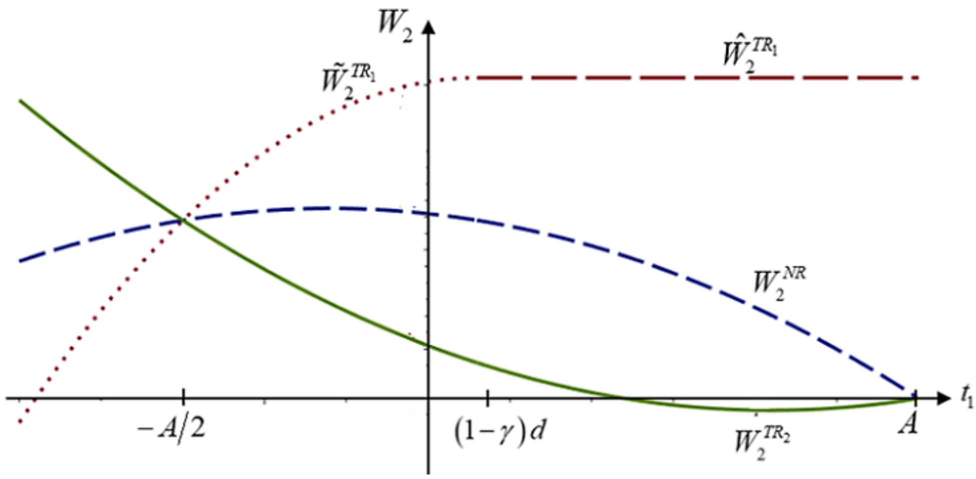
In Figure 1(a), W_1^{NR} is concave and $W_1^{TR_1}$ is convex in t_2 as explained above. For $W_1^{TR_2}$, the unconstrained optimum lies in the segment $t_2 \in (\gamma d, A)$, and is a straight line due to the unique optimal tax $\hat{t}_1^{TR_2} = \gamma d$, whereas the constrained optimum lies in the segment $t_2 \leq \gamma d$ with $\tilde{t}_1^{TR_2} = t_2 - \varepsilon$, and is concave and increasing in t_2.

It is clear that $W_1^{TR_2}$ lies above W_1^{NR} and $W_1^{TR_1}$ in the segment $t_2 \in (-\frac{A}{2}, A)$ and $W_1^{TR_1}$ lies above W_1^{NR} and $W_1^{TR_2}$ in the segment $t_2 < -\frac{A}{2}$. Clearly, for $t_2 \in (\gamma d, A)$, $W_1^{TR_2}$ must dominate all other location equilibria because country 1 has both firms and entirely controls all outputs and hence emissions without any constraint. Obviously, for any $t_2 \leq \gamma d$, undercutting tax t_2 in order to keep all firms also pays country 1 until $t_2 = -\frac{A}{2} + \varepsilon$ is reached at which point net profits (producer surplus plus taxes/minus subsidies) are zero for country

1 in the TR_2-location equilibrium.[11] At subsidy level $t_2 = -\frac{A}{2}$, country 1 is indifferent between keeping its firm (NR-location equilibrium) and its firm leaving (TR_1-location equilibrium) but would certainly not undercut taxes of country 2 as net profits would be negative. Finally, for any subsidy level $t_2 < -\frac{A}{2}$, country 1 prefers if its firm relocates to country 2, again, because its net profits would be negative under the NR-location equilibrium. This suggests that the main driving force for choosing taxes/subsidies are net profits.



(a) Country 1



(b) Country 2

Figure 1: Ranking of the Welfare Levels of Countries without BCAs

From country 1's perspective (Figure 1(a)), starting at tax level $t_2 = \gamma d$, and

[11] We need to keep in mind that $\tilde{W}_1^{TR_2}$ intersects with W_1^{NR} at $t_2 = -\frac{A}{2} + \varepsilon$. As ε is assumed to be close to zero, this detail cannot be seen in Figure 1(a). The same applies for country 2 in Figure 1(b).

gradually lowering tax t_2, implies that the optimal response of country 1 is to marginally undercut taxes of country 2 to keep both firms. For any t_2, matching ($t_1 = t_2$; NR-location equilibrium) instead of undercutting ($t_1 = t_2 - \varepsilon$; TR_2-location equilibrium) constitutes a large and discrete loss of net profits of one firm, a marginal loss of consumer surplus, which cannot be compensated by a marginal gain in the form of reduced damages. Similarly, starting at any tax level $t_2 < A$ and gradually lowering t_2, choosing a higher tax than country 2 ($t_1 > t_2$; TR_1-location equilibrium) instead of matching ($t_1 = t_2$; NR-location equilibrium) constitutes a large discrete loss of net profits of one firm, but has no impact on the consumer surplus and damages. Hence, it is not surprising that there is a race-to-the-bottom equilibrium in the light of global pollution, which we derive in the subsequent subsections.

For country 2, displayed in Figure 1(b), welfare levels and incentives in the three location equilibria are very similar. All welfare functions of the three location equilibria literally intersect at $t_1 = -\frac{A}{2}$. From country 2's perspective, the welfare function of location equilibrium TR_1 is above NR and TR_2 for all $t_1 \in (-\frac{A}{2}, A)$ and, below $t_1 = -\frac{A}{2}$, the welfare function of the TR_2-location equilibrium is above the two other welfare functions. The only difference is that the constrained optimum of the TR_ℓ-equilibrium does not start below tax level γd but below $(1-\gamma)d$, as these are the marginal damages of country 2. Pulling all results together, we can make the following statement.

Lemma 1. The Best Responses of Countries under the No-BCA Regime

i. If $D_i' < t_j < A$, $\hat{W}_i^{TR_\ell} > W_i^{NR} > W_i^{TR_k}$.

ii. If $-\frac{A}{2} < t_j \leq D_i'$, $\tilde{W}_i^{TR_\ell} > W_i^{NR} > W_i^{TR_k}$.

iii. If $t_j \leq -\frac{A}{2}$, $W_i^{TR_k} \geq W_i^{NR} > \tilde{W}_i^{TR_\ell}$.

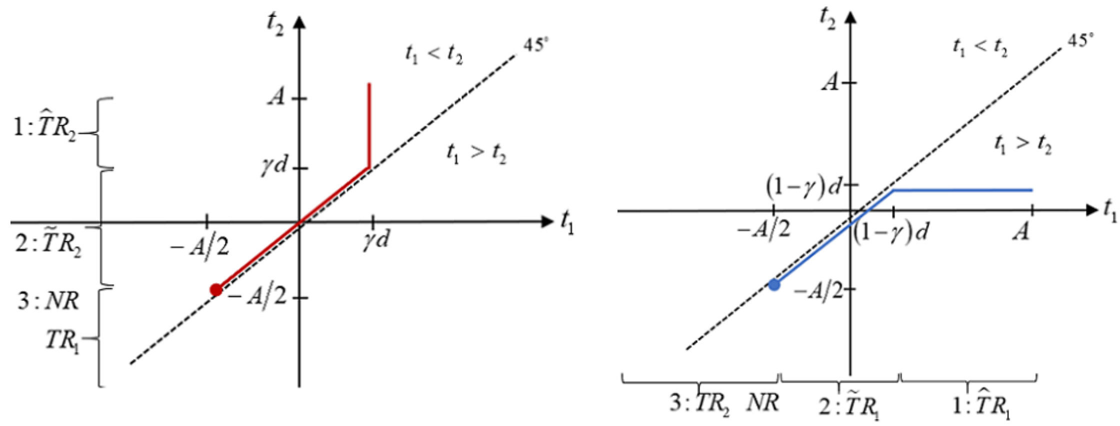
Hence, the best response of country 1 and country 2 are given by:

$$t_1 \begin{cases} = \gamma d & \text{if } \gamma d < t_2 < A \\ = t_2\text{-}\varepsilon & \text{if } -\frac{A}{2} < t_2 \leq \gamma d \\ \geq t_2 & \text{if } t_2 \leq -\frac{A}{2} \end{cases} \tag{13}$$

and

$$t_2 \begin{cases} = (1-\gamma)\,d & \text{if } (1-\gamma)\,d < t_1 < A \\ = t_1\text{-}\varepsilon & \text{if } -\frac{A}{2} < t_1 \leq (1-\gamma)\,d \\ \geq t_1 & \text{if } t_1 \leq -\frac{A}{2} \end{cases} \tag{14}$$

Proof. See Appendix B.2. □



(a) Country 1

(b) Country 2

Figure 2: Best Response Functions of Countries without BCAs

The best response functions are illustrated in Figure 2, each having three segments. The first segment is a vertical (horizontal) line associated with the dominant unconstrained optimum. That is, each country sets its unconstrained carbon tax at marginal damages D_i'. In this range, the location equilibrium is $T\hat{R}_2$ for country 1 and $T\hat{R}_1$ for country 2. The second segment lies marginally above (below) the 45°-line and is upward sloping, which implies that carbon taxes are strategic complements. In this segment, the best response of countries is to undercut each other's taxes, which is associated with the constrained optimum. Hence, the location equilibrium is $T\tilde{R}_2$ for country 1 and $T\tilde{R}_1$ for country 2. The third segment corresponds to any negative tax level equal or below $-\frac{A}{2}$. At $t_j = -\frac{A}{2}$, matching, $t_i = t_j$, as well as choosing any higher tax, $t_i > t_j$ are best responses. For $t_j < -\frac{A}{2}$, the best response of country i is any tax $t_i > t_j$. Thus, in the third segment, the best response is not unique and therefore not drawn. The corresponding location equilibrium for $t_j < -\frac{A}{2}$ is TR_1 for country 1 and TR_2 for country 2, and at $t_j = -\frac{A}{2}$ it is the same plus additionally, NR for both countries.[12]

3.2 Simultaneous Game

In this subsection, we assume countries solve the first stage simultaneously. In Figure 3, we combine the best response functions of both countries to obtain the Nash equilibrium (NE) in pure strategies without BCAs. We denote the Nash

[12]It is worthwhile noting that these best response functions are very different from the single-valued and continuously downward sloping best response functions in models with fixed plant location, which imply that carbon taxes are strategic substitutes and players have unique responses.

equilibrium of this game by $\left(t_1^{NE*}, t_2^{NE*}\right)$.

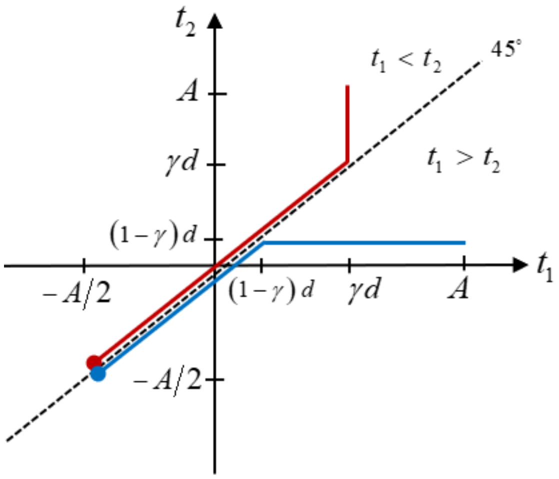


Figure 3: Nash Equilibrium without BCAs

Proposition 2. Nash Equilibrium under the No-BCA Regime

Without BCAs, the pure strategy Nash equilibrium is a subsidy $t_1^{NE*} = t_2^{NE*} = -\frac{A}{2}$*, with location equilibrium* NR*, which constitutes a race-to-the-bottom.*

Proof. Follows directly from Figure 3, noting that $t_1 > t_2 = -\frac{A}{2}$ and $t_2 > t_1 = -\frac{A}{2}$ are not mutually best responses.[13] $\square$

Three features are interesting about Proposition 2. First, both countries are stuck at the bottom, being afraid that their firms relocate their plants abroad, a phenomenon known as 'regulatory chill' (Neumayer, 2001). In a model of fixed plant location, equilibrium taxes would also be below optimal levels due to the incentive to shift rents, but there would be no race-to-the-bottom. Second, the NE is symmetric irrespective of the asymmetry among countries in terms of their damage evaluations. Again, in a model with fixed plant location equilibrium, this would be different as long as marginal damages are different (i.e., $\gamma > 0.5$ in our model). Also, in a model with endogenous plant location but local pollution, an asymmetric NE could occur.[14] Third, in our model, there is only one NE, i.e. the race-to-the-bottom equilibrium. Both countries suffer from damages irrespective

[13] Strictly speaking, all tax levels $t_i \in \left[-\frac{A}{2}, -\frac{A}{2} + \varepsilon\right]$ are Nash equilibria with NR-location equilibrium, where $-\frac{A}{2} + \varepsilon$ is the point at which countries stops undercutting and become indifferent between NR and $\tilde{TR}_\ell$ location equilibria. However, as $\varepsilon \to 0$, this range converges to point $-\frac{A}{2}$, and hence a unique equilibrium. See footnote 11.

[14] For local pollution, an asymmetric NE may emerge with all plants located in the country with lower evaluation of environmental damages. See Rauscher (1995) and Hoel (1997).

of the location of production. Hence, governments undercut each other's carbon tax until the net gains from profits, consumer surplus and tax revenues are exhausted. Our result is qualitatively similar to Rauscher (1995) who considers in his extensions a global pollutant. In his model, countries compete towards a zero tax rate as he does not consider profits in governments welfare function and assumes a monopolistic market structure. In contrast, in a model with local pollution, a second not-in-my-backyard equilibrium could emerge with high carbon taxes such that firms either exit the market or move abroad (Markusen et al., 1995).

3.3 Sequential Game

Suppose country 1 is the leader. Country 1 chooses the point associated with its highest welfare on country 2' best response function. We argue that there are two potential Stackelberg equilibria (SE): 1) SE_1^1 with $t_1^L = t_2^F = -\frac{A}{2}$ and the NR-location equilibrium and 2) SE_1^2 with $t_1^L > t_2^F = (1-\gamma)d$ and the TR_1-location equilibrium with the subscript referring to the leader and the superscript denotes the number of the equilibrium.

From Figure 2(b) and the discussion above, it is clear that any point on segment 3 of country 2's best response function with $t_1 < -\frac{A}{2}$ cannot be attractive to country 1 because country 1 would have to subsidise both firms associated with a loss. Only $t_1 = -\frac{A}{2}$ on segment 3 could potentially qualify as an optimal point. If country 1 sets its tax level at $t_1^L = -\frac{A}{2}$ as a leader, then there are basically two best responses by the follower: $t_1^L = t_2^F = -\frac{A}{2}$ with NR, and $t_1^L = -\frac{A}{2} < t_2^F$ with TR_2. Therefore, we need a tie-breaking rule to have a unique equilibrium. In line with our previous assumption, and, because both countries are indifferent between both equilibria, we assume that firms remain in their home country. Hence, the first Stackelberg equilibrium SE_1^1 is $t_1^L = t_2^F = -\frac{A}{2}$ with NR. This tie-breaking rule also allows for a straightforward comparison between the simultaneous and sequential scenario.

Considering segment 1 and 2 on country 2's best response function, SE_1^2 with $t_1^L > t_2^F = (1-\gamma)d$ and the TR_1-location equilibrium on segment 1 could be another potential equilibrium candidate if country 1 perceives damages to be sufficiently high, i.e., if γd is sufficiently large. Any point on segment 2 can be ruled out because this would lead to the same location equilibrium, but a lower tax level chosen by the follower, i.e., $t_2^F < (1-\gamma)d$, which, as shown in Appendix B.3, leads to a lower welfare level for country 1.

It remains to determine for which conditions SE_1^1 and SE_1^2 will emerge, which is simply tested by the following comparison.

$$W_1^{L*}\left(t_1^{L*} > t_2^{F*} = (1-\gamma)\,d\right) \geq (<) W_1^{L*}\left(t_1^{L*} = t_2^{F*} = -A/2\right) \text{ if } A \leq (>) \hat{A}_1^{TR_1}$$

where $\hat{A}_1^{TR_1} = d\left(2\gamma + \frac{2}{5}\right)$. Note if $A = \hat{A}_1^{TR_1}$, country 1 is indifferent between the two location equilibria, though the follower, country 2, prefers TR_1. Hence, we select SE_1^2 according to the Pareto-criterion.

Proposition 3. Stackelberg Equilibrium under the No-BCA Regime if Country 1 is the Leader

i. The first equilibrium SE_1^1 is a subsidy $t_1^{L*} = t_2^{F*} = -\frac{A}{2}$, with location equilibrium NR if $A > \hat{A}_1^{TR_1}(d,\gamma)$.

ii. The second equilibrium SE_1^2 is a tax $t_1^{L*} > t_2^{F*} = (1-\gamma)\,d$, with location equilibrium TR_1 if $A \leq \hat{A}_1^{TR_1}(d,\gamma)$.

iii. $\hat{A}_1^{TR_1}(d,\gamma)$ increases in γ and d. That is, the higher the damage evaluation of country 1, the more likely it is that the second equilibrium emerges.

Proof. See Appendix B.3. □

The first Stackelberg equilibrium is the same as the race-to-the-bottom Nash equilibrium. The second Stackelberg equilibrium is different, which may be viewed as a chicken equilibrium, with the leader, being the wise chicken, giving in, setting taxes sufficiently high as to avoid the race-to-the-bottom. In both equilibria, profits plus taxes or minus subsidies are zero in country 1. Hence, country 1 chooses between the two equilibria by trading off environmental damages against consumer surplus. Intuitively, the larger γ and d compared to A , the larger are damages in country 1 compared to consumer surplus, and, therefore, the larger will be the threshold level above which country 1 chooses the first race-to-the-bottom equilibrium.

If country 2 is the leader, similar arguments hold, also implying two potential Stackelberg equilibria, SE_2^1 and SE_2^2, with the first one being the race-to-the-bottom equilibrium and the second the chicken equilibrium, also with a threshold such that if $A > \hat{A}_2^{TR_2}$ the first equilibrium emerges and if $A \leq \hat{A}_2^{TR_2}$ the second equilibrium materialises, with $\hat{A}_2^{TR_2} = d\left(2(1-\gamma) + \frac{2}{5}\right)$.

Proposition 4. Stackelberg Equilibrium under the No-BCA Regime if Country 2 is the Leader

i. The first equilibrium SE_2^1 is a subsidy $t_2^{L*} = t_1^{F*} = -\frac{A}{2}$ with location equilibrium NR for all $\gamma \geq 0.7$ or for all $\gamma < 0.7$ if $A > \hat{A}_2^{TR_2}(d, \gamma)$.

ii. The second equilibrium SE_2^2 is a tax $t_2^{L*} > t_1^{F*} = \gamma d$ with the location equilibrium TR_2 for all $\gamma < 0.7$ if $A \leq \hat{A}_2^{TR_2}(d, \gamma)$.

iii. $\hat{A}_2^{TR_2}(d, \gamma)$ increases in d and decreases in γ. That is, the higher the damage evaluation of country 2, the more likely it is that the second equilibrium emerges.

Proof. See Appendix B.3. □

Hence, Propositions 3 and 4 are qualitatively similar. The only difference in Proposition 4 is that for the second equilibrium to emerge, we do not only require $A \leq \hat{A}_2^{TR_2}(d, \gamma)$, but also $\gamma < 0.7$, as otherwise our condition for positive production levels $A > d$ cannot be satisfied as shown in Appendix B.3. It is probably not surprising that $\hat{A}_1^{TR_1}(d, \gamma) \geq \hat{A}_2^{TR_2}(d, \gamma)$ for $\gamma \geq 0.5$, due to the higher damage evaluation of country 1 than country 2. Hence, the range of parameter values for which country 1 is choosing the chicken equilibrium if it is the leader is larger than if country 2 is the leader.

3.4 Comparison under the No-BCA Regime

In this subsection, we briefly compare the Nash equilibrium with the Stackelberg equilibria, and compare those non-cooperative equilibria with the social optimum.

Corollary 1. Comparison of Non-cooperative No-BCA Equilibria and the Social Optimum

i. A sequential instead of a simultaneous choice of taxes is Pareto-improving for both countries and leads to less global emissions if the marginal damages of the Stackelberg leader are sufficiently large.

ii. (a) If $A \leq \hat{A}_2^{TR_2}$ for all $\gamma < 0.7$, each country prefers to be the follower. That is, there is a second mover advantage.

(b) If $\hat{A}_2^{TR_2} < A \leq \hat{A}_1^{TR_1}$, both countries are better off if country 1 acts as the Stackelberg leader.

(c) If $\hat{A}_1^{TR_1} < A$, it makes no difference who is the leader since the outcome is always NR with a race-to-the-bottom.

iii. The non-cooperative carbon tax equilibria in the simultaneous and the sequential game imply a lower effective tax rate, hence, larger global emissions, and lower global welfare than in the social optimum.

iv. The difference between global welfare and global emissions in the social optimum and any of the non-cooperative equilibria increases in the global damage parameter d.

Proof. See Appendix B.4. □

Results i as well as ii can be related to the literature on imperfect competition of firms if choices are strategic complements due to upward sloping best response functions, for instance Gal-Or (1985), assuming symmetric players, shows that there is a second mover advantage and that Stackelberg equilibria are Pareto-improving. In our model, this is relevant in segment 2 and point $t_1 = t_2 = -\frac{A}{2}$ on segment 3, which implies upward sloping best replies. Moreover, players are 'sufficiently symmetric' for $\gamma < 0.7$, but we also need $A \leq \hat{A}_2^{TR_2}$ in our model to have a qualitatively similar result (Corollary 1, i and ii (a)). However, in all other cases, we get different results. If $\hat{A}_2^{TR_2} < A \leq \hat{A}_1^{TR_1}$ (Corollary 1, ii (b)), country 1 prefers to be the leader and if $A > \hat{A}_1^{TR_1}$ (Corollary 1, ii (c)), leadership makes no difference. Thus, in our model, asymmetry and multiple incentives in the objective function (e.g., the trade-off between consumer surplus and environmental damages), give new results.

Result iii is a direct implication of the fact that, in any non-cooperative equilibria, global externalities are not fully internalised. More interesting is result iv; the difference or gap between the social optimum and any of the non-cooperative equilibria increases in the damage parameter d. Noting that a proxy of the net benefits from production and consumption is $A = a - c$ in our model, $A > d$ for an interior solution, an increase of d implies that $\frac{A}{d}$ is decreasing, holding A constant. Hence, the higher global marginal damages compared to the net benefits from production and consumption, the larger would be the gains from full cooperation.

4 Climate Policy Equilibria: BCA Regime

As mentioned in Section 2, we only consider the possibility that country 1 can impose BCAs on imports if $t_1 > t_2$, and consider the possibility that country 2 could do the same if $t_2 > t_1$ in Appendix E. Hence, we have the same location equilibria as under the No-BCA regime, except if $t_1 > t_2$, in which case we have partial instead of total relocation: 1) NR if $t_1 = t_2$, i.e., no relocation of firms; 2) PR_1 if $t_1 > t_2$, i.e., firm 1, with its plant supplying the market in country 2, partially relocates to country 2; 3) TR_2 if $t_1 < t_2$, i.e., all plants are located in country 1. We proceed as in Section 3, namely first derive best responses and then determine Nash and Stackelberg equilibria.

4.1 Best Responses

Since countries have asymmetric incentives, we treat each country separately in this subsection. In order to derive equilibrium taxes, we can exclusively focus on the PR_1-location equilibrium with $t_1 > t_2$, as the other location equilibria are the same as in Section 3. The welfare functions of country 1 and country 2 in the PR_1-location equilibrium are given by:

$$W_1^{PR_1} = CS_1 + \pi_{11} + T_1 - D_1 + BCA_1 \text{ if } t_1 > t_2\,, \tag{15}$$

where $BCA_1 = (t_1 - t_2)\, x_{21}$

$$W_2^{PR_1} = CS_2 + \pi_{12} + \pi_{21} + \pi_{22} + T_2 - D_2 \text{ if } t_1 > t_2\,. \tag{16}$$

Substituting equilibrium output levels given in Appendix C.1 into (15), we obtain the welfare function of country 1:

$$\begin{aligned} W_1^{PR_1} = \frac{1}{2}\left(\frac{2}{3}A - \frac{2}{3}t_1\right)^2 + \frac{(A-t_1)^2}{9} + t_1\left(\frac{1}{3}(A-t_1)\right) \\ (t_1 - t_2)\left(\frac{1}{3}(A-t_1)\right) - \gamma d\left(\frac{4A - 2t_1 - 2t_2}{3}\right) \end{aligned} \tag{17}$$

noting that $W_1^{PR_1}$ is strictly concave in t_1. The optimal tax rate of country 1 follows from

$$\frac{\partial W_1^{PR_1}}{\partial t_1} = \frac{1}{3}(-2t_1 + t_2 + 2\gamma d) = 0 \Leftrightarrow \hat{t}_1^{PR_1}(t_2) = \gamma d + \frac{1}{2}t_2. \tag{18}$$

We call $\hat{t}_1^{PR_1}(t_2)$, as given in (18), the standard best response function of country 1 in the PR_1-location equilibrium, which is upward sloping. That is, for country 1, taxes are strategic complements in this location equilibrium. Since BCA revenues depend on the difference between the two national tax levels, country 1 has an incentive to impose a higher tax if t_2 increases in order to capture more tariff revenues.

$\hat{t}_1^{PR_1}(t_2)$ can be viewed as the unconstrained optimum in the PR_1-location equilibrium. Unfortunately, we need to consider two additional constrained optima. The first emerges from the non-negativity of output levels, with the non-negativity constraint (NNC): $\hat{t}_1^{PR_1} < A$ (see Section 2). If the unconstrained tax would be given by $\hat{t}_1^{PR_1} \geq A$, country 1 deviates from its optimal tax and sets $\ddot{t}_1 = A - \varepsilon$. This follows from the concavity of the welfare of country 1 with respect to t_1 in the PR_1-location equilibrium and the fact that the constraint $\hat{t}_1^{PR_1} < A$ requires a tax rate t_1 below the optimal level. The second constraint is the BCA-constraint, which requires $t_2 < \hat{t}_1^{PR_1}$. If the unconstrained tax would be given by $t_2 \geq \hat{t}_1^{PR_1}$, country 1 deviates by setting $\breve{t}_1 = t_2 + \varepsilon$, which is the lowest possible tax level higher than t_2. Again, this follows from the concavity of the welfare function and the constraint which requires the tax to be above the optimal level.

Substituting equilibrium output levels given in Appendix C.1 into (16), we obtain the welfare function of country 2:

$$\begin{aligned} W_2^{PR_1} &= \frac{1}{2}\left(\frac{2}{3}A - \frac{2}{3}t_2\right)^2 + 2\left(\frac{(A-t_2)^2}{9}\right) + \frac{(A-t_1)^2}{9} \\ &\quad + t_2\left(A - \frac{2}{3}t_2 - \frac{1}{3}t_1\right) - (1-\gamma)\, d\left(\frac{4A - 2t_1 - 2t_2}{3}\right), \end{aligned} \tag{19}$$

noting that $W_2^{PR_1}$ is strictly concave in t_2. The optimal tax rate of country 2 follows from

$$\begin{aligned} \frac{\partial W_2^{PR_1}}{\partial t_2} &= \frac{A}{9} - \frac{4}{9}t_2 - \frac{1}{3}t_1 + \frac{2}{3}(1-\gamma)d = 0 \\ &\Leftrightarrow \hat{t}_2^{PR_1}(t_1) = \frac{1}{4}A + \frac{3}{2}(1-\gamma)\, d - \frac{3}{4}t_1, \end{aligned} \tag{20}$$

where the standard best response function of country 2, $\hat{t}_2^{PR_1}(t_1)$, in the PR_1-location equilibrium, as given in (20), is downward sloping. From country 2's point of view, carbon taxes are strategic substitutes. If t_1 increases, the plant of firm 2, which supplies country 1, faces a higher carbon tariff, which reduces the producer surplus and the tax revenues obtained from this plant in country 2. Therefore, country 2 has an incentive to raise its tax as a countervailing measure. However, such a tax increase negatively affects consumers in country 2 and profits of the other two plants located in country 2 supplying its market. In our model, the second effect dominates the first effect, and, hence, the best response of country 2 is to lower its taxes if country 1's tax increases.

We need also to consider the best response of country 2 if it is not able to choose its unconstrained optimal tax as given in (20). The BCA-constraint requires $t_1 > \hat{t}_2^{PR_1}(t_1)$, which can be shown to imply that $t_1 > \bar{t}_1$ must hold with $\bar{t}_1 = \frac{1}{7}A + \frac{6}{7}(1-\gamma)\,d$. We find that if the BCA-constraint is satisfied, this is sufficient that the non-negativity constraint $\hat{t}_2^{PR_1} < A$ holds. Therefore, if $t_1 > \bar{t}_1$, country 2's best response is to set the unconstrained tax $\hat{t}_2^{PR_1}$. However, if $t_1 \leq \bar{t}_1$, country 2 will deviate from its optimal tax and will choose its constrained tax $\tilde{t}_2 = t_1 - \varepsilon$.

Pulling results together, our just obtained results for the new PR_1-location equilibrium as well as our previous results obtained for the NR- and TR_2-location equilibria in Section 3, we obtain:

Welfare Levels of Country 1 for Three Location Equilibria:

1. Case $t_1 < t_2$: $\hat{W}_1^{TR_2}$ and $\tilde{W}_1^{TR_2}$ as defined under the No-BCA regime.
2. Case $t_1 = t_2$: W_1^{NR} as defined under the No-BCA regime.
3. Case $t_1 > t_2$

   (a) Let $\hat{W}_1^{PR_1}$ be $W_1^{PR_1}\left(t_1 = \hat{t}_1^{PR_1}(t_2)\right)$ as a function of t_2, in the range $t_2 < 2\gamma d$ and $t_2 < 2A - 2\gamma d$. Country 1 chooses its unconstrained tax $\hat{t}_1^{PR_1}(t_2)$ as given in (18).

   (b) Let $\ddot{W}_1^{PR_1}$ be $W_1^{PR_1}\left(t_1 = \ddot{t}_1 = A - \varepsilon\right)$ as a function of t_2, in the range $t_2 \geq 2A - 2\gamma d$. Country 1 chooses its constrained tax $\ddot{t}_1$, which is marginally smaller than the maximum feasible level.

   (c) Let $\breve{W}_1^{PR_1}$ be $W_1^{PR_1}\left(t_1 = \breve{t}_1 = t_2 + \varepsilon\right)$ as a function of t_2, in the range $t_2 \geq 2\gamma d$. Country 1 chooses its constrained tax at $\breve{t}_1$, which is marginally above the lowest possible tax rate and hence marginally above t_2.

Welfare Levels of Country 2 for Three Location Equilibria:

1. Case $t_1 > t_2$

   (a) Let $\hat{W}_2^{PR_1}$ be $W_2^{PR_1}\left(t_2 = \hat{t}_2^{PR_1}(t_1)\right)$ as a function of t_1, in the range $t_1 > \bar{t}_1$. Country 2 chooses its unconstrained tax level $\hat{t}_2^{PR_1}(t_1)$ as given in (20).

   (b) Let $\tilde{W}_2^{PR_1}$ be $W_2^{PR_1}\left(t_2 = \tilde{t}_2 = t_1 - \varepsilon\right)$ as a function of t_1, in the range $t_1 \leq \bar{t}_1$. Country 2 chooses its constrained tax rate $\tilde{t}_2$ marginally below country 1's tax rate t_1.
2. Case $t_2 = t_1$: W_2^{NR} as defined under the No-BCA regime.

3. Case $t_1 < t_2$: $W_2^{TR_2}$ as defined under the No-BCA regime.

The details of the welfare functions defined above are given in Appendix C.1.

In order to derive the best response function of country 1 and 2, it is helpful to consider Figures 4 and 5.[15] Let us start to consider country 1's incentives in Figure 4.

The welfare function of country 1 in the TR_2- and NR-location equilibrium are the same as in Figure 1(a). The different in Figure 4 is that the welfare function in the TR_1-location equilibrium has been replaced by the PR_1-location equilibrium where the latter equilibrium is much more attractive than the former to country 1. Country 1 can keep one plant of its own firm, and, most importantly, receives tariff revenues from BCAs. Compared to the NR-location equilibrium, country 1 has both lower profits (since it has one plant fewer) and lower consumer surplus, but also lower damages and positive tariff revenues in the PR_1-location equilibrium. Thus, there exist tax levels t_2 above $-\frac{A}{2}$ for which $\hat{W}_1^{PR_1} > W_1^{NR}$ is possible. Similarly, compared to the TR_2-location equilibrium, in the PR_1-location equilibrium, country 1 has three plants fewer and hence less profits, as well as lower consumer surplus, but also lower damages and positive tariff revenues. The intersection point of $\tilde{W}_1^{TR_2}$ and $\hat{W}_1^{PR_1}$ is at $\underline{t_2}$, which is larger than $t_2 = -\frac{A}{2}$.

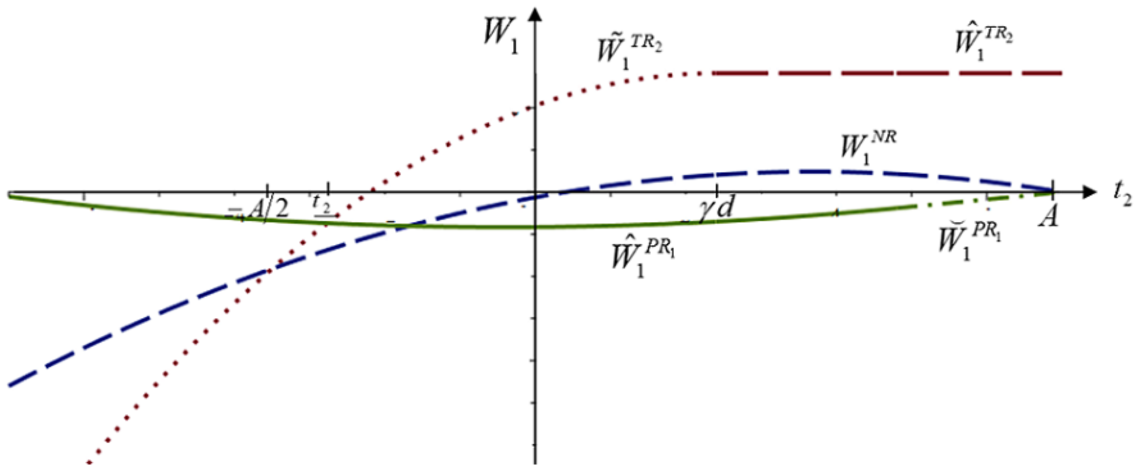


Figure 4: Ranking of the Welfare Levels of Country 1 with BCAs

From Figure 4, it is evident that there are two relevant segments when it comes to ranking welfare: i) any tax $t_2 > \underline{t_2}$ implies that TR_2 and ii) any tax $t_2 < \underline{t_2}$ implies

[15]The Welfare functions in Figure 4 correspond to a particular case, which is called case $(2\,(a))$ in Lemma 2, and which is provided in Appendix C.2. For other cases, there will no qualitative changes.

that PR_1 is the preferred location equilibrium. Pulling all features together, leads to Lemma 2.

Lemma 2. The Best Response of Country 1 under the BCA Regime

i. If $\gamma d < t_2 < A$, $\hat{W}_1^{TR_2} > W_1^{PR_1}$ *and* $\hat{W}_1^{TR_2} > W_1^{NR}$.

ii. If $\underline{t_2} < t_2 \leq \gamma d$, $\tilde{W}_1^{TR_2} > W_1^{PR_1}$ *and* $\tilde{W}_1^{TR_2} > W_1^{NR}$.

iii. If $-\frac{A}{2} < t_2 \leq \underline{t_2}$, $\hat{W}_1^{PR_1} \geq \tilde{W}_1^{TR_2}$ *and* $\hat{W}_1^{PR_1} > W_1^{NR}$.

iv. If $t_2 \leq -\frac{A}{2}$, $\hat{W}_1^{PR_1} > W_1^{NR} > \tilde{W}_1^{TR_2}$.

Hence, the best response of country 1 is given by:

$$t_1 \begin{cases} = \gamma d & \text{if } \gamma d < t_2 < A \\ = t_2\text{-}\varepsilon & \text{if } \underline{t_2} \leq t_2 \leq \gamma d \\ = \hat{t}_1^{PR_1}(t_2) > t_2 & \text{if } t_2 \leq \underline{t_2} \end{cases} \tag{21}$$

Proof. See Appendix C.2.

□

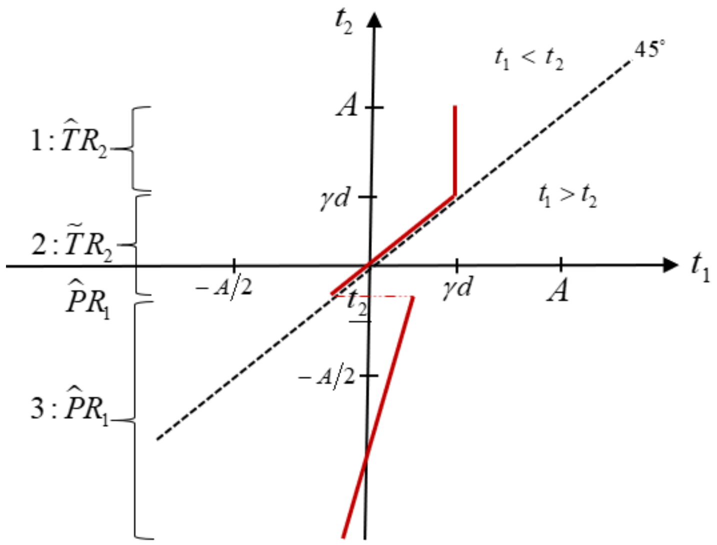


Figure 5: Best Response Function of Country 1 with BCAs

The best response function of country 1 has three segments, like under the No-BCA regime (see Figure 2(a)). The first two segments are similar: a vertical line, i.e., the dominant strategy with the unconstrained TR_2-location equilibrium (denoted by $T\hat{R}_2$), and the marginally undercutting tax segment with the constrained TR_2-location equilibrium (denoted by $T\tilde{R}_2$), which lies slightly above the 45°-line. However, the second segment stops at $\underline{t_2}$ in Figure 5 and not at $t_2 = -\frac{A}{2}$

as in Figure 2(a). Also, different from Figure 2(a), the third part in Figure 5 is upward sloping, which is the standard best reply function in the PR_1-location equilibrium (denoted by PR_1). Segments 2 and 3 have one point in common at $t_2 = \underline{t_2}$. Hence, the best response function is discontinuous at $t_2 = \underline{t_2}$: country 1 is indifferent between the TR_2- and the PR_1-location equilibrium. We note that matching taxes with the NR-location equilibrium is never a best response for any tax level t_2.

Consider now country 2. In Figure 6, for all $t_1 > \bar{t}_1$, country 2 achieves its highest welfare level if it sets its optimal tax level, $\hat{t}_2^{PR_1}$, attracting three plants. However, since $\hat{t}_2^{PR_1}$ is a function of t_1, we notice that $\hat{W}_2^{PR_1}$ is a convex function. For any $t_1 \leq \bar{t}_1$, country 2 responds by choosing its constrained carbon tax in the PR_1-location equilibrium with $\tilde{t}_2 = t_1 - \varepsilon$, and hence $\tilde{W}_2^{PR_1}$ is a strictly concave function. As we explained in the No-BCA regime, undercutting will stop when profits minus subsidies become zero, i.e., marginally above $t_1 = -\frac{A}{2}$.[16] That is, further undercutting means too much subsidies that exceed profits of firms, and, hence, this it is not a best response.

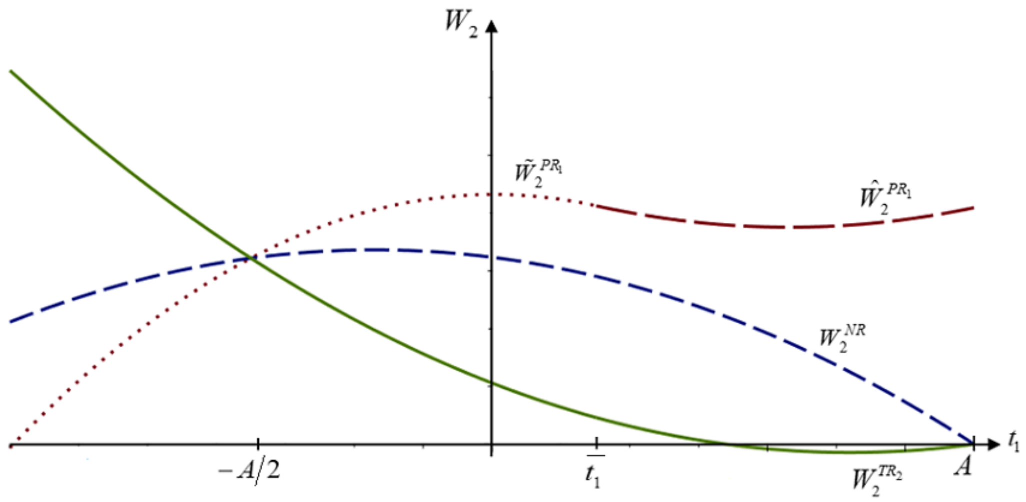


Figure 6: Ranking of the Welfare Levels of Country 2 with BCAs

From the previous analysis, the best response of country 2 is summarised in the following Lemma.

[16] See footnote 11.

Lemma 3. The Best Response of Country 2 under the BCA Regime

i. If $\bar{t}_1 < t_1 < A$, $\hat{W}_2^{PR_1} > W_2^{NR} > W_2^{TR_2}$.

ii. If $-\frac{A}{2} < t_1 \leq \bar{t}_1$, $\tilde{W}_2^{PR_1} > W_2^{NR} > W_2^{TR_2}$.

iii. If $t_1 \leq -\frac{A}{2}$, $W_2^{TR_2} \geq W_2^{NR} > \tilde{W}_2^{PR_1}$.

Hence, the best response of country 2 is given by:

$$t_2 \begin{cases} = \hat{t}_2^{PR_1}(t_1) & \text{if } \bar{t}_1 < t_1 < A \\ = t_1\text{-}\varepsilon & \text{if } -\frac{A}{2} < t_1 \leq \bar{t}_1 \\ \geq t_1 & \text{if } t_1 \leq -\frac{A}{2} \end{cases} \tag{22}$$

Proof. See Appendix C.3. □

The best response function of country 2 in Figure 7 is nearly the same as under the No-BCA regime in Figure 2(b). The difference is that the boundary between segment 1 and 2 is not $t_1 = (1-\gamma)d$ but $t_1 = \bar{t}_1$ and segment 1 is not a horizontal line but a downward sloping function.

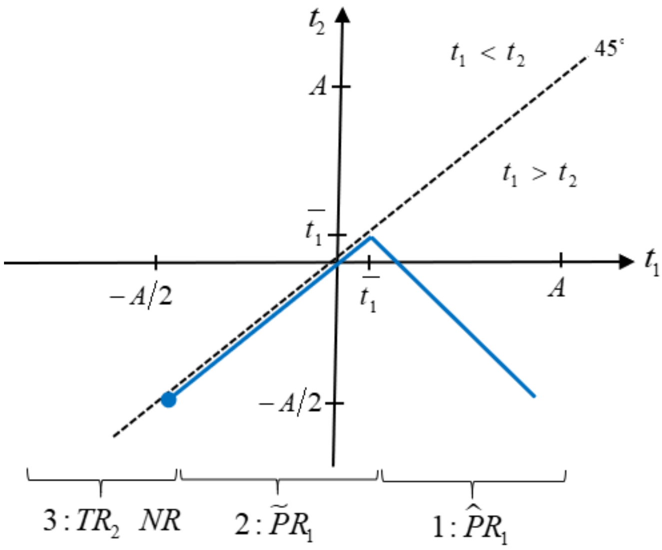
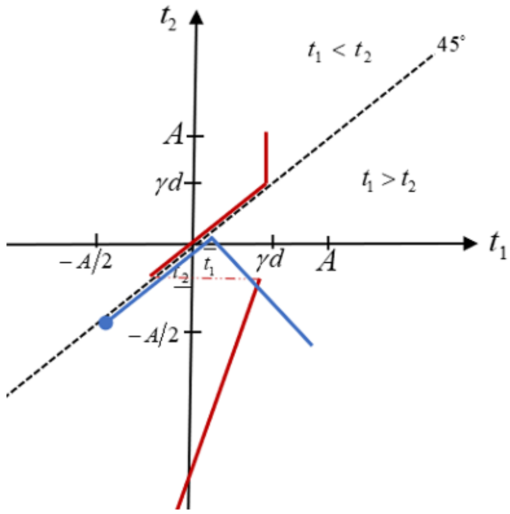


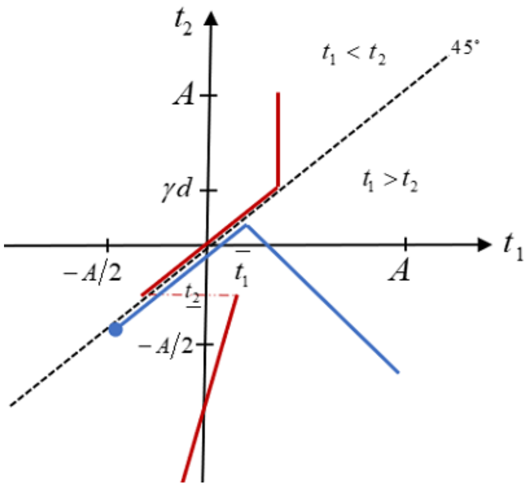
Figure 7: Best Response Function of Country 2 with BCAs

4.2 Simultaneous Game

The fact that country 1 stops undercutting before country 2 under the BCA regime implies that a race-to-the-bottom is not a Nash Equilibrium (NE). Moreover, matching taxes is never a best response for country 1 with BCAs. Hence, the NR-location equilibrium cannot be a NE. We can also rule out the constrained PR_1-location equilibrium ($P\tilde{R}_{1)}$ and the constrained TR_2-location equilibrium ($T\tilde{R}_2$) with mutually undercutting taxes along the 45°-line, as best response functions do not intersect. Therefore, if a NE exits, it must be that country 1 sets a higher tax than country 2 and firm 1 partially relocates with one plant to country 2. This is the unconstrained PR_1-location equilibrium ($P\hat{R}_1$). The equilibrium tax levels of country 1 and 2, $t_1^{NE*(P\hat{R}_1)}$ and $t_2^{NE*(P\hat{R}_1)}$, follow from solving (18) and (20) simultaneously (see Appendix C.4).



(a) Existence of a NE



(b) Non-existence of a NE

Figure 8: Nash Equilibrium with BCAs

In Figure 8(a), the two 'standard best response functions' in the PR_1-location equilibrium intersect and hence a NE exists. However, this may not always be the case, and, hence, a NE may not exist as shown in Figure 8(b). As mentioned by Anderson (1987) and Mintz and Tulkens (1986), this may arise due to the discontinuity and multivaluedness of best response functions. At $\underline{t_2}$, country 1 has two best responses, and there is a discrete jump from one to another segment of the best response function.

For a pure strategy NE to exist, the intersection of the best reply functions in the PR_1-location equilibrium must occur in the range where country 1's best response is located on its standard best response function. That is, we must have $\underline{t_2} \geq t_2^{NE*(P\hat{R}_1)}$. However, if $\underline{t_2} < t_2^{NE*(P\hat{R}_1)}$, the best response of country 1 would be to marginally undercut country 2's tax t_2. Hence, best reply functions in the unconstrained PR_1-location equilibrium do not intersect and a pure strategy NE does not exist.

Proposition 5. Nash Equilibrium under the BCA Regime

i. A pure strategy NE may not exist.

ii. A pure strategy NE exists if $A \leq \overline{A}_{NE}(d, \gamma)$ *for all* $\gamma > 0.855$ *with* $\overline{A}_{NE}(d, \gamma)$ *increasing in* d *and* γ*. That is, a NE exists if global marginal damages and the asymmetry among countries are sufficiently large.*

iii. If a NE exists, $t_1^{NE*}(\hat{PR_1}) > t_2^{NE*}(\hat{PR_1})$, $t_1^{NE*}(\hat{PR_1}) > 0$ *and* $-\frac{A}{2} < t_2^{NE*}(\hat{PR_1}) < 0$*, and the location equilibrium of firms is* PR_1.

Proof. See Appendix C.4, including the precise definition of $\overline{A}_{NE}(d, \gamma)$. □

Thus, if the global damage evaluation parameter d is large enough, a NE may exist under the BCA regime. We recall that the difference between the NE under the No-BCA regime and the social optimum in terms of global emissions and global welfare increases in the parameter d (see Corollary 1, in particular part iv). Hence, if d is large, the potential gains from cooperation would be large. It is exactly under those conditions where BCAs are an effective policy measure and help to escape the race-to-the-bottom, though socially optimal levels cannot be obtained, as we know from Proposition 1.

4.3 Sequential Game

We first solve for equilibrium taxes under the BCA regime if country 1 is the leader. Country 1 can choose any point on country 2' best response function as shown in Figure 7. There are basically three options.

- Option 1: On segment 3, without BCAs, $t_1 = t_2 = -\frac{A}{2}$ is the only point that could be attractive to country 1 with the NR-location equilibrium because for both countries subsidies above $-\frac{A}{2}$ are never attractive.
- Option 2: On segment 1, with BCAs, country 1 can set $t_1^L > \bar{t}_1$ to induce relocation of one plant of its home firm, which results in the unconstrained PR_1-location equilibrium. Equilibrium tax of country 1 follows from maximising $W_1^{PR_1}\left(t_1, \hat{t}_2^{PR_1}(t_1)\right)$ with respect to t_1, taking into account the best response $\hat{t}_2^{PR_1}(t_1)$ of country 2, as given by (20). The corresponding equilibrium tax of country 2 follows by substituting optimal t_1 into (20). We denote equilibrium taxes by $t_1^{L*}(\hat{PR_1})$ and $t_2^{F*}(\hat{PR_1})$.
- Option 3: On segment 2, with BCAs, country 1 sets a tax $-\frac{A}{2} < t_1^L \leq \bar{t}_1$ which would be undercut by country 2, $t_2 = t_1^L - \varepsilon$, which results in the constrained PR_1-location equilibrium.

It turns out that country 1 prefers options 1 and 2 over option 3. Whether option 1 or 2 is chosen depends on parameter values.[17]

Proposition 6. Stackelberg Equilibrium under the BCA Regime if Country 1 is the Leader

i. The first equilibrium SE_1^1 is a subsidy $t_1^{L*} = t_2^{F*} = -\frac{A}{2}$, with location equilibrium NR if $A > \overline{A}_1^{PR_1}(d, \gamma)$.

ii. The second equilibrium SE_1^2 is $t_1^{L*}(\hat{PR_1}) > t_2^{F*}(\hat{PR})$, where *$t_1^{L*}(\hat{PR_1}) > 0$ and $t_2^{F*}(\hat{PR}) > -\frac{A}{2}$, with location equilibrium PR_1 if $A < \overline{A}_1^{PR_1}(d, \gamma)$.*

iii. $\overline{A}_1^{PR_1}(d, \gamma)$ increases in d and γ. That is, the higher the damage evaluation of country 1, the more likely it is that the second equilibrium emerges.

Proof. See Appendix C.5, including the precise definition of $\overline{A}_1^{PR_1}(d, \gamma)$. □

There exists a threshold $\overline{A}_1^{PR_1}(d, \gamma)$ below which country 1 would prefer partial relocation of its home firm.[18] As under the No-BCA regime, this equilibrium occurs if marginal damages of country 1, γd, are sufficiently large compared to A, a proxy for the net gains from production and consumption. This second equilibrium allows the leader to escape the race-to-the-bottom, with higher effective taxes in both countries, hence lower global emissions. Different from the No-BCA regime, the leader does not have to choose between the NR- and the TR_1-location equilibrium, but between the NR- and the PR_1-location equilibrium where the latter choice is more attractive. Hence, the chicken equilibrium is more attractive to country 1 under the BCA regime. This explains why the threshold under the BCA regime, $\overline{A}_1^{PR_1}(d, \gamma)$, is higher than the threshold under the No-BCA regime, $\hat{A}_1^{TR_1}(d, \gamma)$, i.e., $\hat{A}_1^{TR_1} < \overline{A}_1^{PR_1}$. In other words, for a given value of A, it requires lower marginal damages in country 1 under the BCA than the No-BCA regime in order for country 1 to opt for escaping the race-to-the-bottom.

We now consider that country 2 is the Stackelberg leader, even though one may argue that this possibility is not very likely, given that only country 1 can implement BCAs. Never, we consider this possibility for completeness.

It turns out that the derivation of Stackelberg equilibria is much more complicated due to the jump of the best response function of country 1 under this regime.

[17] In principle, the leader, country 1, could also decide not to use BCAs, set $t_1^L > (1-\gamma)\,d$, in order to induce the total relocation of its firm, resulting in the TR_1-location equilibrium. In Appendix C.5, we show that this option is dominated by option 2.

[18] If $A = \overline{A}_1^{PR_1}$, country 1 is indifferent between the two location equilibria. Hence, according to the Pareto-criterion, we select SE_1^2 for all $\gamma < 0.98$, while SE_1^1 for all $\gamma \geq 0.98$.

Therefore, we just mention a couple of observations with reference to Figure 5, while the details are provided in Appendix C.6.

First, the NR-location equilibrium is no longer a choice for country 2 as it is not part of country 1's best response function under the BCA regime. Second, for the TR_2-location equilibrium, country 2 could choose either $t_2^L > \gamma d$ to induce the unconstrained $\hat{TR_2}$-location equilibrium or $\underline{t_2} < t_2^L \leq \gamma d$ to induce the constrained $\tilde{TR_2}$-location equilibrium. Third, for the PR_1-location equilibrium, country 2 could choose $t_2^L < \underline{t_2}$ by maximising $W_2^{PR_1}\left(t_2, \hat{t}_1^{PR_1}(t_2)\right)$ with respect to t_2, taking country 1's best response function in the PR_1-location equilibrium into account as given in (18) such that the unconstrained $\hat{PR_1}$-location equilibrium emerges with $t_2^{L*}(\hat{PR_1}) < t_1^{F*}(\hat{PR_1})$. However, if $t_2^{L*}(\hat{PR_1}) > \underline{t_2}$, country 2 has to deviate from its unconstrained tax level and set $t_2^{L*} \lesssim \underline{t_2}$ which follows from the concavity of the welfare of country 2 in the PR_1-location equilibrium. Fourth, if country 2 chooses $t_2 = \underline{t_2}$, country 1 would be indifferent between choosing its optimal tax $\hat{t}_1^{PR_1}$ in the PR_1-location equilibrium or undercutting tax t_2, corresponding to its constrained TR_2-location equilibrium ($\tilde{TR_2}$). Therefore, if country 2 is better off under PR_1, it will set a tax level marginally below $\underline{t_2}$, i.e., $t_2^{L*} \lesssim \underline{t_2}$, that induces country 1 to react on its standard best reply function in the $\hat{PR_1}$-location equilibrium. While if country 2 prefers the $\tilde{TR_2}$-location equilibrium, it will set a tax level marginally above $\underline{t_2}$, i.e., $t_2^{L*} \gtrsim \underline{t_2}$, such that country 1 responds by undercutting this tax level, $t_1 = t_2 - \varepsilon$.[19]

As Proposition 7 spells out, country 2 will either choose the PR_1-location equilibrium or the constrained TR_2-location equilibrium around $\underline{t_2}$ on country 1's best response function, which gives rise to two Stackelberg equilibria.

Proposition 7. Stackelberg Equilibrium under the BCA Regime if Country 2 is the Leader

i. The first equilibrium SE_2^1 is: (a) $t_2^{L*}(\hat{PR_1}) < t_1^{F*}(\hat{PR_1})$ if $A < \overline{A}_2^{PR_1}$ for all $\gamma \geq \hat{\gamma} \approx 0.881$ or (b) $t_2^{L*} \lesssim \underline{t_2} < \hat{t}_1^{PR_1} = t_1^{F*}$ if $\overline{A}_2^{PR_1} \leq A \leq \overline{A}_2^{TR_2}$ with $t_1^{F*}(\hat{PR_1}) > 0$; $-\frac{A}{2} < t_2^{L*} < 0$. The equilibrium location of firms is PR_1.

ii. The second equilibrium SE_2^2 is $t_2^{L*} \gtrsim \underline{t_2}$, and $t_1^{F*} = t_2^{L*} - \varepsilon$, with $-\frac{A}{2} <$ *$t_2^{L*}, t_1^{F*} < 0$. The equilibrium location of firms is TR_2, if $A > \overline{A}_2^{TR_2}(d, \gamma)$.*

[19] For a similar case, see Dixit (1979), though in a different context. Dixit explains that if the leader chooses its strategy at the point where the follower is indifferent between two best responses, there is technically no equilibrium if the leader prefers one response by the follower to the other. In such a case, the leader can choose its strategy slightly larger or smaller than this indifference point such that the follower responds in a way that gives the leader the highest payoff.

iii. $\overline{A}_2^{TR_2}(d,\gamma)$ *increases in* d *and decreases in* γ.

Proof. See Appendix C.6, including the precise definitions of $\overline{A}_2^{PR_1}$ and $\overline{A}_2^{TR_2}$. □

The incentives of the leader are very different for country 2 than country 1. Whereas country 1 can choose between the race-to-the-bottom equilibrium with no relocation and the chicken PR_1-location equilibrium if its damages are sufficiently high, country 2 faces the choice between partial relocation of firm 1 under the BCA regime (PR_1-location equilibrium) and total relocation of its firm to country 1 without BCAs (TR_2-location equilibrium). Both equilibria imply lower global emissions than in the race-to-the-bottom equilibrium. So the 'chicken choice' for environmental reasons is in the range $A \leq \overline{A}_2^{TR_2}$. For $A > \overline{A}_2^{TR_2}$, country 2 would need to provide large subsidies to three plants of which one faces tariffs on exports to country 1 (and hence net profits are low) in the PR_1-location equilibrium, i.e., $t_2 \approx \underline{t_2}$ is negative and large in absolute terms, which does not pay. Hence, total relocation is preferred and no BCAs are imposed.

4.4 Comparison under the BCA Regime

In this subsection, we compare the Nash equilibrium with the Stackelberg equilibria under the BCA regime, and all non-cooperative equilibria with the social optimum. In order to compare the three non-cooperative equilibria, we need to assume that a Nash equilibrium exists. It turns out that if the sufficient conditions for the existence of a NE hold, implying an unconstrained PR_1-location equilibrium, the same is true for the Stackelberg equilibria regardless whether country 1 and 2 is the leader. More specifically, if country 1 is the leader, the second equilibrium SE_1^2 with taxes $t_1^{*L}(\hat{PR_1}) > t_2^{*F}(\hat{PR_1})$ emerges (see Proposition 6) and if country 2 is the leader, the first equilibrium SE_2^1 with (a) $t_2^{*L}(\hat{PR_1}) < t_1^{*F}(\hat{PR_1})$ will materialize (see Proposition 7). In other words, the chicken equilibria emerge under Stackelberg leadership, implying that global marginal damages d are sufficiently large compared to the benefits from production and consumption, approximated by our parameter $A = a - c$.

Since already in the Nash equilibrium, a race-to-the-bottom is avoided under the BCA regime, it cannot be expected that Stackelberg leadership is always Pareto-improving and leads to lower global emissions compared to the Nash equilibrium as this was the case under the No-BCA regime (see Corollary 1). Some interesting findings are summarized below.

Corollary 2. Comparison of Non-cooperative BCA Equilibria and the Social Optimum

Suppose the sufficient conditions for the existence of a Nash equilibrium under the BCA regime hold.

i. If country 1 is the Stackelberg leader, global welfare and global emissions are higher than in the Nash equilibrium. Stackelberg leadership is Pareto-improving if $\gamma \geq \check{\gamma} \approx 0.94$.

ii. If country 2 is the Stackelberg leader, global welfare is higher than in the Nash equilibrium, Stackelberg leadership is Pareto-improving and leads to lower global emissions if $\gamma < \ddot{\ddot{\gamma}} \approx 0.98$.

iii. Global emissions are higher and global welfare is higher under Stackelberg leadership of country 1 than under Stackelberg leadership of country 2.

iv. Global emissions in all three non-cooperative equilibria are higher than in the social optimum.

Under those conditions for which a Nash equilibrium does not exist, global emissions under Stackelberg leadership may be lower than in the social optimum. This is the case if global marginal damages d are low compared to the gains from production and consumption A, i.e., if the gains from cooperation are small.

Proof. See Appendix C.7. □

Corollary 2 confirms that Stackelberg leadership is not always Pareto-improving and even global welfare may not increase under Stackelberg leadership of country 2. However, Stackelberg leadership always increases global welfare if country 1 is the leader and if country 2 is the leader if asymmetries are not too pronounced. Under leadership of country 2, a global welfare increase goes along with lower global emissions, but under leadership of country 1 the reverse is true, global emissions increase. Taken result ii and iii together, Stackelberg leadership always increases global welfare compared to the Nash equilibrium if a social planner could choose who is the Stackelberg leader, depending on the value of γ.

The very last result is in line with Hoel (1997) who shows that under endogenous plant location, equilibrium policy levels can lead to an overshooting in that equilibrium global emissions are below socially optimal levels. However, in our model, this only occurs if the gains from cooperation are small. In any case, global welfare levels are always below the socially optimal level, as follows directly from Proposition 1.

5 Comparison of Climate Policies across Regimes: The Role of BCAs

In this section, we compare equilibria with and without BCAs in order to analyse the effect of BCAs. A sensible comparison must assume the same sequence of moves under both regimes.

5.1 Simultaneous Game

In order to evaluate the effects of BCAs if countries choose their taxes simultaneously, we must assume that a NE exists under the BCA regime. According to Proposition 5, this implies $A \leq \overline{A}_{NE}(d, \gamma)$ and $\gamma > 0.855$. Based on Proposition 2 and 5, we can draw the following conclusions.

Corollary 3. The Effect of BCAs in the Simultaneous Game

i. BCAs lead to higher taxes in both countries, which leads to lower global emissions and higher global welfare.

ii. BCAs always increase the welfare of country 1 and increase the welfare of country 2 if $\gamma \leq \bar{\gamma} \approx 0.96$

iii. BCAs change the equilibrium location of firms from NR *to* PR_1*.*

Proof. See Appendix D.1. □

Lower global emissions follows directly from the fact that, under the BCA regime, countries avoid the race-to-the-bottom equilibrium, and, instead, the PR_1-location equilibrium with higher taxes in both countries emerges. This is welfare improving for both countries, except for country 2 if its damage evaluation, $(1-\gamma)d$, is very low. However, even if country 2 would be worse off, aggregate welfare is higher under the BCA regime. Clearly, BCAs must improve country 1's welfare position as BCAs equip country 1 with an additional policy tool.

Thus, under a simultaneous move scenario, BCAs achieve for what they are designed: they protect country 1, which is more environmentally conscious, and allow to implement a more ambitious climate policy which also raises global welfare. Only if country 2 gives hardly any weight to environmental damages are BCAs not Pareto-improving, a qualification which is hardly surprising. We may also recall that a NE under the BCA regime exists if the gains from cooperation would be large. So BCAs are useful when it matters.

5.2 Sequential Game

In order to analyse the effect of BCAs in the sequential game, we consider three parameter ranges when country 1 is the leader and the same is true when country 2 is the leader, though ranges under different leadership do not coincide as Corollary 4 spells out.

Corollary 4. The Effect of BCAs in the Sequential Game

Let E, W, W_1, W_2 denote global emissions, global welfare and individual welfare of country 1 and 2, respectively, and let $\uparrow$ denote an increase and $\downarrow$ denote a decrease of a variable. Let the parameter ranges be defined as in Propositions 3, 4, 6 and 7.

(1) Suppose country 1 is the Stackelberg leader. BCAs imply the following changes compared to No-BCAs:

i. Region F, $A \leq \hat{A}_1^{TR_1}$: $TR_1 \Longrightarrow PR_1$; $E \downarrow$; $W_1 \uparrow$, $W_2 \downarrow$, $W \uparrow$ if $\gamma \leq \tilde{\gamma} \approx 0.83$.

ii. Region G, $\hat{A}_1^{TR_1} < A \leq \overline{A}_1^{PR_1}$: $NR \Longrightarrow PR_1$; $E \downarrow$; $W_1 \uparrow$, $W_2 \uparrow$ if $\gamma < \breve{\gamma} \approx 0.98$, $W \uparrow$ if $\gamma < \breve{\gamma} \approx 0.98$.

iii. Region H, $\overline{A}_1^{PR_1} \leq A$: $NR \Longrightarrow NR$; no changes.

(2) Suppose country 2 is the Stackelberg leader. BCAs imply the following changes compared to No-BCAs

i. Region M, $A \leq \hat{A}_2^{TR_1}$ and $\gamma < 0.7$: $TR_2 \Longrightarrow (PR_1)$ TR_2; ($E \uparrow$; $W_1 \downarrow$, $W_2 \uparrow$, $W \downarrow$) no changes.

ii. Region N, $\hat{A}_2^{TR_1} < A \leq \bar{A}_2^{TR_2}$: $NR \Longrightarrow PR_1$; $E \downarrow$; $W_1 \uparrow$, $W_2 \uparrow\downarrow$ if $\gamma < \mathring{\gamma} \approx 0.96$, $W \uparrow$.

iii. Region O, $\bar{A}_1^{TR_2} < A$: $NR \Longrightarrow \tilde{TR}_2$; $E \downarrow$; $W_1 \uparrow$, $W_2 \downarrow$, $W \uparrow$ if $A < \underline{A}^O$.

Proof. See Appendix D.2, including the precise definition of $\underline{A}^O$.[20] □

Let us start by considering the effects if country 1 is the leader. Without BCAs, countries escape the race-to-the-bottom in region F, while in regions G and H, this is not the case. With BCAs, in region F and G, the PR_1-location equilibrium emerges, whereas in region H, the NR-location equilibrium with the race-to-the-bottom emerges. Thus, in region H, BCAs make no difference. However, this

[20]If $A = \overline{A}_1^{PR_1}$ and $\gamma < 0.98$, country 1, as a leader, chooses PR_1, and, hence this threshold level is included in region G. Whereas if $\gamma \geq 0.98$, country 1 chooses NR, and, hence this threshold level is included in region H. See footnote 18.

region is where parameter A is relatively large compared to d (recall: $\overline{A}_1^{PR_1}(d,\gamma)$ increases in d) and hence the potential gains from cooperation would anyway be small.

In region G, BCAs help to avoid the race-to-the-bottom, and hence global emissions are lower than without BCAs. Both countries benefit from BCAs, except if country 2 with the lower damage evaluation hardly values damages at all. This result seems intuitive as well as the fact that country 1 is always better off. BCAs provide country 1 with an additional policy tool (and country 1 is the Stackelberg leader), which must be advantageous and at worst neutral.

This is very similar in region F, even though without BCAs the race-to-the-bottom is avoided. For country 1, clearly, the PR_1-location equilibrium is much better than the TR_1-location equilibrium and not surprisingly, for country 2 this is reversed, as the TR_1-location equilibrium provides the highest welfare to country 2. Therefore, even though global emissions decrease under the BCA regime, global welfare only increases provided a certain threshold of γ is not exceeded. That is, country 2's damage evaluation is not too low.

Thus, taken together, BCAs lower global emissions whenever the potential gains from cooperation would be large. They normally raise global welfare if the environmental damage evaluation in country 2 is not too low, which, in our model, means that the damage evaluation in both countries is not too different. If the race-to-the-bottom equilibrium is already avoided without BCAs, then BCAs only benefit country 1.

We now consider the effects if country 2 is the leader. We argued above that this constellation is not very likely as only country 1 can implement BCAs, however, we consider this possibility for completeness.

Given that country 1 decides about the implementation of BCAs, country 1 would never implement BCAs in region M, as without BCAs, it has all firms and can choose its optimal tax. In regions N and O, BCAs improve upon the race-to-the-bottom equilibrium, which emerges without BCAs. Global emissions are always lower, global welfare is always higher in region N and in most cases in region O. However, although country 1 is always better off with BCAs, this is not always true for country 2. Similar to above, we can conclude that BCAs help the country which implements this policy, country 1 in our model, lowers global emissions and in most cases increases global welfare, even though the country on which BCAs are imposed may be worse off.

6 Conclusions

One main obstacle of more ambitious policies to address global warming are leakage effects. One particular form of leakage is the relocation of the production of firms, which may even imply that firms close down and move abroad if environmental regulation increases their production cost too much. This is in particular relevant for emission-intensive industries which trade internationally. In order to capture this phenomenon and the discussion surrounding it, we set up an intra-industry trade model and studied an emission tax competition game between two asymmetric countries when both, the location choice of firms and the policy choice of governments, are endogenous. Asymmetry implied in our model that countries evaluate the damages from global pollution differently. We solved a three-stage game in which governments first choose their emission tax, then firms choose the location of their two plants (one producing for the home and one for the foreign market), and finally firms choose their production levels. We considered two policy regimes. Under the No-BCA regime, each government imposes a carbon tax on the production within its national boundaries. This regime served as a benchmark to study the effect of border carbon adjustments, abbreviated BCAs, which we called the BCA regime. Under this regime, the government which sets a higher carbon tax can additionally impose a tariff on imports, which have been produced facing a lower tax abroad. BCAs fully adjust the difference between the carbon taxes in the two countries. This implies that all plants supplying the country that imposes BCAs face the same effective carbon tax. As under the BCA regime a Nash equilibrium (i.e., a simultaneous choices of taxes) in the tax game may not exist (due to discontinuity of reaction functions), we also considered Stackelberg equilibria (i.e., sequential choices of taxes).

Without BCAs, the effective taxes that firms face, and hence their profits, are based on the location of production. Thus, each firm will locate with its two plants in the country which sets a lower carbon tax. Thus, a government setting a higher carbon tax than its rival government will see its firm relocating abroad. If countries choose their climate policies simultaneously, this leads to a fierce tax competition in which each government has an incentive to undercut the other government's carbon tax in order to keep their firm or even attract the foreign firm. That is, we showed that the Nash equilibrium is a 'race to the bottom' leading to symmetric low taxes (or symmetric high subsidies) with high global emissions, irrespective of the absolute value of environmental damages and the degree of asymmetry among countries in terms of the evaluation of damages. In this equilibrium, each firm remains with its two plants in the country of origin, which we called no relocation (NR). Even if the environmentally more friendly government recognises damages as being important for the welfare of its country, it is rational for governments to lower taxes if the foreign government does so as well. By lowering taxes gradually, environmental damages increase marginally, but the gain in terms of profits is discrete, (i.e., avoiding the loss of profits because the own firm does not locate abroad and/or increasing profits because the foreign firm is attracted to the home country). This is the leakage dilemma of environmentally concerned governments. Interestingly, in a Stackelberg equilibrium with a sequential tax choice, governments may be able to avoid the race-to-the-bottom, and instead a 'wise chicken' equilibrium emerges. This is the case if the Stackelberg leader values environmental damages sufficiently high, recognises the disastrous outcome of the race-to-the-bottom equilibrium and hence sets a high carbon tax such that its firm relocates to the follower's country. This total relocation (TR) equilibrium leads to less global emissions and, more importantly, is Pareto-improving for both countries compared to the Nash equilibrium.

Also under the BCA regime, the pressure on the race-to-the-bottom was reduced even if governments choose their taxes simultaneously. A tariff allows the environmentally more concerned government to set a higher tax without the danger of loosing both plants. BCAs create an equal playing field for all production sold to the home market. Thus, higher taxes do not lead to total relocation (TR) but only to partial relocation (PR). Moreover, the government imposing carbon tariffs at the border gains a strategic advantage because it is able to shift tax revenues from abroad to home. In the simultaneous game, a Nash equilibrium may not exist (due to non-continuous reaction functions), though we showed that it exists when the potential gains from cooperation would be large. That is, it exists if global marginal damages are sufficiently large. Also in the sequential game, the wise chicken equilibrium, associated with lower global emissions, is more likely to emerge than the race-to-the-bottom equilibrium. Hence, taken together, our results showed that BCAs can support more ambitious climate policies and can be globally welfare improving under the more general assumption of endogenous plant location. However, we also demonstrated that BCAs will always fall short of achieving the socially optimal global welfare level.

In this paper, we considered a “weak form” of BCAs (also sometimes referred to as partial BCAs), which was a carbon tariff. However, also a “strong form ” (also sometimes called full BCAs) has been suggested where tariffs are complemented by export rebates. That is, emission taxes on exports are reduced to the lower foreign emission tax level. This creates not only an equal playing field for goods sold to the market in which the higher emission tax is levied on production, but also for goods sold to the market in which the lower tax is imposed on production. De facto, this means to move from a production- to consumption-based emission tax. It may also imply that BCAs do not only avoid total relocation of firms located in environmentally friendly countries abroad, but may even avoid partial relocation. Hence, it would be interesting to explore whether "strong BCAs" could further improve global welfare. This is not obvious in a strategic context as export rebates per se increase pollution. Another extension could be to analyse whether and under which conditions BCAs are an effective policy tool to enforce a socially optimal solution. We showed that the social optimum can be obtained under three location equilibria, which greatly differ in the individual welfare levels of countries. In addition, BCAs result in partial relocation of firms from the more to the less environmentally concerned country. Hence, it is not straightforward to predict a priori whether BCAs would be a credible threat to enforce a socially optimal solution in this set-up.

Acknowledgement

We would like to thank Dr. Javier Rivas for his helpful comments. We are also grateful for the comments of the attendees of the PhD Brown Bag Seminar at the University of Bath, and the feedback of the participants of EAERE 2019 in Manchester, in particular, Prof. Knut Einar Rosendahl, and the participants of the third AERNA Workshop on Game Theory and the Environment 2019 in Valencia, Spain.

References

Anderson, S. (1987). Spatial competition and price leadership. *International Journal of Industrial Organization*, 5(4):369–398.

Anouliés, L. (2015). The strategic and effective dimensions of the border tax adjustment. *Journal of Public Economic Theory*, 17(6):824–847.

Baksi, S. and Chaudhuri, A. R. (2017). International trade and environmental cooperation among heterogeneous countries. In Kayalıca, M. Ö., Çağatay, S., and Mıhçı, H., eds., *Economics of International Environmental Agreements*, pp. 97–115. Routledge.

Barnett, A. H. (1980). The Pigouvian tax rule under monopoly. *The American Economic Review*, 70(5):1037–1041.

Barrett, S. (1994). Strategic environmental policy and intrenational trade. *Journal of Public Economics*, 54(3):325–338.

Böhringer, C., Balistreri, E. J., and Rutherford, T. F. (2012). The role of border carbon adjustment in unilateral climate policy: overview of an energy modeling forum study EMF 29. *Energy Economics*, 34:S97–S110.

Brander, J. A. and Spencer, B. J. (1985). Export subsidies and international market share rivalry. *Journal of International Economics*, 18(1-2):83–100.

Branger, F. and Quirion, P. (2014). Would border carbon adjustments prevent carbon leakage and heavy industry competitiveness losses? Insights from a meta-analysis of recent economic studies. *Ecological Economics*, 99:29–39.

Conrad, K. (1993). Taxes and subsidies for pollution-intensive industries as trade policy. *Journal of Environmental Economics and Management*, 25(2):121–135.

Copeland, B. R. (1996). Pollution content tariffs, environmental rent shifting, and the control of cross-border pollution. *Journal of International Economics*, 40(3-4):459–476.

De Santis, R. A. and Stähler, F. (2009). Foreign direct investment and environmental taxes. *German Economic Review*, 10(1):115–135.

Dijkstra, B. R., Mathew, A. J., and Mukherjee, A. (2011). Environmental regulation: an incentive for foreign direct investment. *Review of International Economics*, 19(3):568–578.

Dixit, A. (1979). A model of duopoly suggesting a theory of entry barriers. *The Bell Journal of Economics*, pp. 20–32.

Eerola, E. (2006). International trade agreements, environmental policy, and relocation of production. *Resource and Energy Economics*, 28(4):333–350.

Eskeland, G. S. and Harrison, A. E. (2003). Moving to greener pastures? Multinationals and the pollution haven hypothesis. *Journal of Development Economics*, 70(1):1–23.

Eyland, T. and Zaccour, G. (2014). Carbon tariffs and cooperative outcomes. *Energy Policy*, 65:718–728.

Fischer, C. and Fox, A. K. (2012). Comparing policies to combat emissions leakage: border carbon adjustments versus rebates. *Journal of Environmental Economics and Management*, 64(2):199–216.

Fredriksson, P. G., List, J. A., and Millimet, D. L. (2003). Bureaucratic corruption, environmental policy and inbound US FDI: theory and evidence. *Journal of Public Economics*, 87(7-8):1407–1430.

Gal-Or, E. (1985). First mover and second mover advantages. *International Economic Review*, 26(3):649–653.

Hecht, M. and Peters, W. (2018). Border adjustments supplementing nationally determined carbon pricing. *Environmental and Resource Economics*, 73(1):93–109.

Helm, D., Hepburn, C., and Ruta, G. (2012). Trade, climate change, and the political game theory of border carbon adjustments. *Oxford Review of Economic Policy*, 28(2):368–394.

Hoel, M. (1996). Should a carbon tax be differentiated across sectors? *Journal of Public Economics*, 59(1):17–32.

Hoel, M. (1997). Environmental policy with endogenous plant locations. *The Scandinavian Journal of Economics*, 99(2):241–259.

Ikefuji, M., Itaya, J.-i., and Okamura, M. (2016). Optimal emission tax with endogenous location choice of duopolistic firms. *Environmental and Resource Economics*, 65(2):463–485.

Janeba, E. (1998). Tax competition in imperfectly competitive markets. *Journal of International Economics*, 44(1):135–153.

Kellenberg, D. K. (2009). An empirical investigation of the pollution haven effect with strategic environment and trade policy. *Journal of International Economics*, 78(2):242–255.

Kennedy, P. W. (1994). Equilibrium pollution taxes in open economies with imperfect competition. *Journal of Environmental Economics and Management*, 27(1):49–63.

Manderson, E. and Kneller, R. (2012). Environmental regulations, outward FDI and heterogeneous firms: are countries used as pollution havens? *Environmental and Resource Economics*, 51(3):317–352.

Markusen, J. R. (1975). International externalities and optimal tax structures. *Journal of International Economics*, 5(1):15–29.

Markusen, J. R., Morey, E. R., and Olewiler, N. (1995). Competition in regional environmental policies when plant locations are endogenous. *Journal of Public Economics*, 56(1):55–77.

Markusen, J. R., Morey, E. R., and Olewiler, N. D. (1993). Environmental policy when market structure and plant locations are endogenous. *Journal of Environmental Economics and Management*, 24(1):69–86.

Mintz, J. and Tulkens, H. (1986). Commodity tax competition between member states of a federation: equilibrium and efficiency. *Journal of Public Economics*, 29(2):133–172.

Motta, M. and Thisse, J.-F. (1994). Does environmental dumping lead to delocation? *European Economic Review*, 38(3):563–576.

Neumayer, E. (2001). Do countries fail to raise environmental standards? An evaluation of policy options addressing "regulatory chill". *International Journal of Sustainable Development*, 4(3):231–244.

Petrakis, E. and Xepapadeas, A. (2003). Location decisions of a polluting firm and the time consistency of environmental policy. *Resource and Energy Economics*, 25(2):197–214.

Rauscher, M. (1995). Environmental regulation and the location of polluting industries. *International Tax and Public Finance*, 2(2):229–244.

Sanna-Randaccio, F., Sestini, R., and Tarola, O. (2017). Unilateral climate policy and foreign direct investment with firm and country heterogeneity. *Environmental and Resource Economics*, 67(2):379–401.

Stiglitz, J. (2006). A new agenda for global warming. *The Economists' Voice*, 3(7).

Ulph, A. and Valentini, L. (2001). Is environmental dumping greater when plants are footloose? *The Scandinavian Journal of Economics*, 103(4):673–688.

Wooders, P., Cosbey, A., and Stephenson, J. (2009). Border carbon adjustment and free allowances: responding to competitiveness and leakage concerns. *Round Table on Sustainable Development*, OECD.

Xing, Y. and Kolstad, C. D. (2002). Do lax environmental regulations attract foreign investment? *Environmental and Resource Economics*, 21(1):1–22.

Appendix

A Possible Location Equilibria

Consider the No-BCA regime. Firm k will base its decision where to locate its plant supplying market i on $\Delta\pi_{ki} = \pi_{ki}(i,\ell) - \pi_{ki}(j,\ell)$, $\ell = i, j$. For a given location of firm ℓ´s plant supplying market i, firm k will locate in country i if $\Delta\pi_{ki} > 0$ and will locate in country j if $\Delta\pi_{ki} < 0$.

i) Let $\ell = i$. Then, $\Delta\pi_{ki} = \pi_{ki}(i,i) - \pi_{ki}(j,i) = -\frac{4}{9}\left(t_i - t_j\right)\left(A - t_j\right)$ as $\pi_{ki}\left(i,i\right) = \frac{(A-t_i)^2}{9}$ and $\pi_{ki}\left(j,i\right) = \frac{(A-2t_j+t_i)^2}{9}$, making use of (3) and (4) in the text, noting that the first payoff assumes tax vector (t_i, t_i) and the second (t_j, t_i).

ii) Let $\ell = j$. Then, $\Delta\pi_{ki} = \pi_{ki}(i,j) - \pi_{ki}(j,j) = -\frac{4}{9}\left(t_i - t_j\right)\left(A - t_i\right)$ as $\pi_{ki}\left(i,j\right) = \frac{(A-2t_i+t_j)^2}{9}$ and $\pi_{ki}\left(j,j\right) = \frac{(A-t_j)^2}{9}$, again, using (3) and (4) in the text, and noting that the first payoff assumes tax vector (t_i, t_j) and the second (t_j, t_j). Since $A - t_j > 0$ and $A - t_i > 0$ by assumption, firm k's plant supplying market i has a dominant strategy, irrespective where firm ℓ´s plant supplying market i locates.

It moves its plant to country j if $t_i > t_j$ and moves it to country i if the reverse is true, i.e., $t_i < t_j$. By assumption, if $t_i = t_j$, firm k´s plant supplying market i remains in its home country.

The same considerations apply for firm k´s plant supplying market j. Accordingly, all plants of firm k are based in the same country and hence there are three location equilibria: NR, TR_1 and TR_2. Under the BCA regime, assuming that country i can impose BCAs, it is clear that the same computations apply for production sold to market j. For the protected market i, it is easily checked that $\Delta\pi_{ki} = 0$, irrespective where the competitor plant of firm ℓ supplying market i locates. Consequently, firm k with its plant supplying market i remains it its country of origin. If we let $i = 1$, then NR, PR_1 and TR_2 are possible location equilibria and if $i = 2$, then NR, PR_2 and TR_1.

B Appendix of Section 3

B.1 Welfare Functions of Countries under Possible Location Equilibria

Recall that we assume $A > t_i$ and $A > d$ throughout the paper to ensure positive production levels in all possible location equilibria.

In case $t_i < t_j$, all plants are located in country i and face an effective tax $t_{ki} = t_{kj} = t_i$ $\forall k = 1, 2$, and $i, j = 1, 2$ and $i \neq j$. Inserting these tax levels into (3) in the text, gives the equilibrium output levels under TR_ℓ: $x^*_{ki} = x^*_{kj} = (A - t_i)/3$.

Substituting these equilibrium output levels into (8a) in the text gives the welfare function of country i under TR_ℓ:

$$W_i^{TR_\ell} = \frac{1}{2}\left(\frac{2}{3}A - \frac{2}{3}t_i\right)^2 + 4\left(\frac{A - t_i}{3}\right)^2 + t_i\left(\frac{4}{3}(A - t_i)\right) - D_i^{'}\left(\frac{4}{3}(A - t_i)\right), \quad \text{(A.1)}$$

where $W_i^{TR_\ell}$ is concave in t_i, $\frac{\partial^2 W_i^{TR_\ell}}{\partial t_i^2} = -\frac{4}{3} < 0$. (The four terms are consumer surplus, producer surplus, tax revenues and damages.) Simplification of (A.1) leads to (9) in the text.

$\hat{W}_i^{TR_\ell}$ and $\tilde{W}_i^{TR_\ell}$ are obtained by inserting $\hat{t}_i^{TR_\ell} = D_i^{'}$ and $\tilde{t}_i = t_j - \varepsilon$, respectively, into (A.1):

$$\hat{W}_i^{TR_\ell} = \frac{1}{2}\left(\frac{2}{3}A - \frac{2}{3}D_i^{'}\right)^2 + 4\left(\frac{1}{3}A - \frac{1}{3}D_i^{'}\right)^2, \quad \text{(A.2)}$$

$$\tilde{W}_i^{TR_\ell} = \frac{1}{2}\left(\frac{2}{3}A - \frac{2}{3}(t_j - \varepsilon)\right)^2 + 4\left(\frac{A-(t_j-\varepsilon)}{3}\right)^2 + (t_j - \varepsilon)\left(\frac{4}{3}(A-(t_j-\varepsilon))\right) - D_i^{'}\left(\frac{4}{3}(A-(t_j-\varepsilon))\right), \tag{A.3}$$

where $\tilde{W}_i^{TR_\ell}$ is concave in t_j, $\frac{\partial^2 \tilde{W}_i^{TR_\ell}}{\partial t_j^2} = -\frac{4}{3}$.

In case $t_i = t_j$, each firm locates with its two plants in the country of origin and faces an effective tax $t_{ki} = t_{kj} = t_i = t_j$, which gives equilibrium output levels: $x_{ki}^* = x_{kj}^* = (A - t_i)/3 = (A - t_j)/3$ $\forall k = 1, 2$, and $i, j = 1, 2$ and $i \neq j$.

Inserting these equilibrium output levels in (8b) in the text, setting $t_i = t_j$, gives:

$$W_i^{NR} = \frac{1}{2}\left(\frac{2}{3}A - \frac{2}{3}t_i\right)^2 + 2\left(\frac{A-t_i}{3}\right)^2 + t_i\left(\frac{2}{3}(A-t_i)\right) - D_i^{'}\left(\frac{4}{3}(A-t_i)\right), \tag{A.4}$$

where $\frac{\partial^2 W_i^{NR}}{\partial t_i^2} = -\frac{4}{9} < 0$. Simplification of (A.4) gives (10) in the text.

In case $t_i > t_j$, all plants are located in country j, and face an effective tax $t_{ki} = t_{kj} = t_j$. Inserting these tax levels into (3), gives the equilibrium output levels under TR_k as $x_{ki}^* = x_{kj}^* = (A - t_j)/3$.

In this case, using (8c) in the text, $W_i^{TR_k}$ is given by:

$$W_i^{TR_k} = \frac{1}{2}\left(\frac{2}{3}A - \frac{2}{3}t_j\right)^2 - D_i^{'}\left(\frac{4}{3}(A-t_j)\right), \tag{A.5}$$

which is convex in t_j.

B.2 Proof of Lemma 1

We use Appendix B.1. First, it is clear that each country achieves the highest welfare level if it attracts all plants and imposes its unconstrained carbon tax. Therefore, we always have $\hat{W}_i^{TR_\ell} > W_i^{NR}$ and $\hat{W}_i^{TR_\ell} > W_i^{TR_k}$. Second, W_i^{NR} and $W_i^{TR_k}$ intersect at two levels: $t_j = -\frac{A}{2}$ and $t_j = A$, where $W_i^{NR} \geq W_i^{TR_k}$ for all $t_j \in \left[-\frac{A}{2}, A\right)$ and $W_i^{NR} < W_i^{TR_k}$ for all $t_j < -\frac{A}{2}$. Therefore, for $t_j \in \left(D_i^{'}, A\right)$, we have $\hat{W}_i^{TR_\ell} > W_i^{NR} > W_i^{TR_k}$. Third, $\tilde{W}_i^{TR_\ell}$ intersects with W_i^{NR} at two levels of t_j which are A and $-\frac{A}{2}$ for $\varepsilon \to 0$ where $\tilde{W}_i^{TR_\ell} > W_i^{NR}$ for all $t_j \in \left(-\frac{A}{2}, A\right)$ and $W_i^{NR} > \tilde{W}_i^{TR_\ell}$ for all $t_j \leq -\frac{A}{2}$. Therefore, for $t_j \in \left(-\frac{A}{2}, D_i^{'}\right]$, we have $\tilde{W}_i^{TR_\ell} > W_i^{NR} > W_i^{TR_k}$. Fourth, the three curves intersect at $t_j = -\frac{A}{2}$ for $\varepsilon \to 0$. Nevertheless, strictly speaking, $\tilde{W}_i^{TR_\ell}$ intersects with W_i^{NR} at $t_j = -\frac{A}{2} + \varepsilon$, and at $t_j = -\frac{A}{2}$, $\tilde{W}_i^{TR_\ell} < W_i^{NR} = W_i^{TR_k}$. Finally, for all $t_j < -\frac{A}{2}$, we have $W_i^{TR_k} > W_i^{NR} > \tilde{W}_i^{TR_\ell}$.

B.3 Proof of Propositions 3 and 4

First, we compare the equilibrium welfare levels (denoted by an asterisk henceforth) of each leading country under the two choices. For country 1, we find that $W_1^{L*(T\hat{R}_1)}\left(t_1^{L*} > t_2^{F*} = (1-\gamma)\,d\right) \geq W_1^{L*(NR)}\left(t_1^{L*} = t_2^{F*} = -\frac{A}{2}\right)$ if $A \leq \hat{A}_1^{TR_1} = 2\gamma d + \frac{2}{5}d$. With respect to country 2, we find that $W_2^{L*(T\hat{R}_2)}\left(t_2^{L*} > t_1^{F*} = \gamma d\right) \geq W_2^{L*(NR)}\left(t_2^{L*} = t_1^{F*} = -\frac{A}{2}\right)$ if $A \leq \hat{A}_2^{TR_2} = 2(1-\gamma)d + \frac{2}{5}d$. However, given our condition, $A > d$, we can only have $A \leq \hat{A}_2^{TR_2}$ if $\gamma < 0.7$.

Second, we need to check that if the leader i chooses to let its firm to relocate with both plants, whether it always prefers the follower j to choose its optimal unconstrained tax $\hat{t}_j = D_j^{'}$ instead of undercutting, $t_j = t_i - \varepsilon$. We know that $W_i^{TR_k}$ is strictly convex with minimum $t_j^{min} = A - 3D_i^{'}$. Moreover, W_i^{NR} and $W_i^{TR_k}$ intersect at $t_j = -\frac{A}{2}$, as shown in Appendix B.2. We know that $-\frac{A}{2} < D_j^{'}$, which gives rise to three possibilities.

1) $t_j^{min} \leq -\frac{A}{2} < D_j^{'}$. $\hat{t}_j = D_j^{'}$ is at the upward sloping part of $W_i^{TR_k}$ and undercutting cannot pay.

2) $-\frac{A}{2} < t_j^{min} \leq D_j^{'}$. Again, $\hat{t}_j = D_j^{'}$ is at the upward sloping part of $W_i^{TR_k}$. If $W_i^{TR_k}(\hat{t}_j = D_j^{'}) \geq W_i^{NR}(t_j = -\frac{A}{2})$ undercutting cannot pay and if $W_i^{TR_k}(\hat{t}_j = D_j^{'}) < W(t_j = -\frac{A}{2})$, country i would anyway choose the NR-location equilibrium instead of the TR_k-location equilibrium.

3) $-\frac{A}{2} < D_j^{'} \leq t_j^{min}$. $\hat{t}_j = D_j^{'}$ is at the downward sloping part of $W_i^{TR_k}$ and hence country i would choose the NR-location equilibrium instead of the TR_k-location equilibrium.

B.4 Proof of Corollary 1

(i) The SE improves upon the NE outcome only if the Stackelberg leader chooses the second equilibrium as the first equilibrium is exactly the same as the NE. In the second equilibrium, the leader must be better off than in the first equilibrium and for the follower the same is true because its best location equilibrium materialises. The tax in the second equilibrium is $t_j = D_j^{'} > -\frac{A}{2}$ and hence global emissions must be lower than in the first equilibrium.

(ii) (a) If $A \leq \hat{A}_2^{TR_2}$ for all $\gamma < 0.7$, each country as a leader will choose the TR_k-location equilibrium (SE_i^2 in Proposition 3 and Proposition 4) given that $\hat{A}_2^{TR_2} \leq$

$\hat{A}_1^{TR_1}$ for all $\gamma \geq 0.5$. For the follower, this is the $T\hat{R}_\ell$-location equilibrium which is the best location equilibrium for each country. Hence, each country prefers to be the follower.

(b) If $\hat{A}_2^{TR_2} < A \leq \hat{A}_1^{TR_1}$, country 2 as a leader will choose the NR-location equilibrium (SE_2^1 in Proposition 4), while country 1 chooses the TR_1-location equilibrium over the NR-location equilibrium (SE_1^2 in Proposition 3). Hence, both countries prefer country 1 to be the leader.

(c) Finally, if $\hat{A}_1^{TR_1} < A$, both countries as a leader will choose the NR-location equilibrium, hence it makes no difference who is the leader.

(*iii*) We have $t_S^* - t^{NE*} = \frac{3}{2}d > 0$. If country 1 is the leader, we have $t_S^* \leq t_2^{F*} = (1-\gamma)d$ if $A \geq d(2\gamma+1)$ in the second equilibrium which requires $A \leq \hat{A}_1^{TR_1}$. We find $d(2\gamma+1) > \hat{A}_1^{TR_1}$, and, hence $t_S^* > t_2^{F*}$. Similarly, if country 2 leads, in the second equilibrium which requires $A \leq \hat{A}_2^{TR_2}$, $t_S^* \leq t_1^{F*} = \gamma d$ if $A \geq d(3-2\gamma)$. However, $d(3-2\gamma) > \hat{A}_2^{TR_2}$, thus, $t_S^* > t_1^{F*}$.

(*iv*) Global welfare and global emission levels in the social optimum are $W_S^* = (A-d)^2$ and $E_S^* = 2(A-d)$, respectively, while under the non-cooperative equilibria: $W^{NE*} = A(A-2d)$ and $E^{NE*} = 2A$ in the simultaneous game, and $W^{SE_1^2*} = \frac{4}{9}(A-(1-\gamma)d)(2(A-d)-\gamma d)$, $E^{SE_1^2*} = \frac{4}{3}(A-(1-\gamma)d)$ and $W^{SE_2^2*} = \frac{4}{9}(A-\gamma d)(2A-3d+\gamma d)$, $E^{SE_2^2*} = \frac{4}{3}(A-\gamma d)$ in the sequential game if country 1 and country 2 is the leader, respectively. $\frac{\partial W_S^* - W^{NE*}}{\partial d} = 2d > 0$, $\frac{\partial W_S^* - W^{SE_1^2*}}{\partial d} > 0$ if $A < d(2\gamma+1)$, which must hold as long as $A \leq \hat{A}_1^{TR_1}$, and $\frac{\partial W_S^* - W^{SE_2^2*}}{\partial d} > 0$ if $A < d(3-2\gamma)$, which is larger than $\hat{A}_2^{TR_2}$. Note that the first equilibrium in the sequential game is exactly the same as the NE. In addition, $\frac{\partial E^{NE*} - E_S^*}{\partial d} = 2$, $\frac{\partial E^{SE_1^2*} - E_S^*}{\partial d} = \frac{2}{3}(1+2\gamma)$ and $\frac{\partial E^{SE_2^2*} - E_S^*}{\partial d} > 0$ for all $\gamma < 1.5$, which holds as we assume $\gamma \leq 1$.

C Appendix of Section 4

C.1 Welfare Functions of Countries under Possible Location Equilibria

In the case $t_1 > t_2$, there is a partial relocation of firm 1. Both plants supplying country 1 face $t_{k1} = t_1 \; \forall k = 1,2$, while those plants supplying country 2 face $t_{k2} = t_2 \; \forall k = 1,2$. Hence, inserting these tax levels into (3) gives the equilibrium output levels under PR_1: $x_{k1}^* = (A-t_1)/3$ and $x_{k2}^* = (A-t_2)/3$ for $k = 1,2$.

We start first with the welfare functions of country 1 as follows:

(a) Inserting $\hat{t}_1^{PR_1}(t_2) = \gamma d + \frac{1}{2}t_2$ from (18) into (17), we obtain:

$$\begin{aligned}\hat{W}_1^{PR_1} =& \frac{1}{2}\left(\frac{2}{3}A - \frac{1}{3}t_2 - \frac{2}{3}\gamma d\right)^2 + \frac{(2A - 2\gamma d - t_2)^2}{36} \\ &+ \frac{(2\gamma d - t_2)(2A - 2\gamma d - t_2)}{12} + \frac{(2\gamma d + t_2)(2A - 2\gamma d - t_2)}{12} \\ &- \gamma d\left(\frac{4A - 2\gamma d}{3} - t_2\right),\end{aligned} \tag{A.6}$$

where $\frac{\partial^2 \hat{W}_1^{PR_1}}{\partial t_2^2} = \frac{1}{6}$, i.e., $\hat{W}_1^{PR_1}$ is convex in t_2.

(b) Inserting $\ddot{t}_1 = A - \varepsilon$ into (17) gives:

$$\ddot{W}_1^{PR_1} = \frac{(2A - t_2 - \varepsilon)\varepsilon}{3} - \gamma d\left(\frac{2(A - t_2 + \varepsilon)}{3}\right), \tag{A.7}$$

which is linear in t_2.

(c) Inserting $t_1 = \breve{t}_1 = t_2 + \varepsilon$ into (17) gives:

$$\begin{aligned}\breve{W}_1^{PR_1} =& \frac{1}{2}\left(\frac{2}{3}A - \frac{2}{3}(t_2 + \varepsilon)\right)^2 + \frac{(A - (t_2 + \varepsilon))^2}{9} + \varepsilon\left(\frac{A - (t_2 + \varepsilon)}{3}\right) \\ &(t_2 + \varepsilon)\left(\frac{A - (t_2 + \varepsilon)}{3}\right) - \gamma d\left(\frac{4A - 4t_2 - 2\varepsilon}{3}\right),\end{aligned} \tag{A.8}$$

which can be shown to be linear in t_2.

The remaining cases are the same as in Appendix B.1.

For country 2, the welfare functions under the PR_1-location equilibrium are given by:

(a) Inserting $\hat{t}_2^{PR_1}(t_1) = \frac{1}{4}A + \frac{3}{2}(1-\gamma)d - \frac{3}{4}t_1$ from (20) into (19) delivers:

$$\begin{aligned}\hat{W}_2^{PR_1} =& \frac{1}{2}\left(\frac{1}{2}A + \frac{1}{2}t_1 - (1-\gamma)d\right)^2 + \frac{(A + t_1 - (2(1-\gamma)d))^2}{8} \\ &+ \frac{(A - t_1)^2}{9} + \frac{(A - 3t_1 + 6(1-\gamma)d)(5A + t_1 - 6(1-\gamma)d)}{24} \\ &- (1-\gamma)d\left(\frac{7}{6}A - \frac{1}{6}t_1 - (1-\gamma)d\right),\end{aligned} \tag{A.9}$$

where $\frac{\partial^2 \hat{W}_2^{PR_1}}{\partial t_1^2} = \frac{17}{36}$, i.e., $\hat{W}_2^{PR_1}$ is convex in t_1.

(b) Inserting $\tilde{t}_2 = t_1 - \varepsilon$ into (19) gives:

$$\tilde{W}_2^{PR_1} = \frac{1}{2}\left(\frac{2}{3}A - \frac{2}{3}(t_1 - \varepsilon)\right)^2 + 2\left(\frac{(A-(t_1-\varepsilon))^2}{9}\right) + \frac{(A-t_1)^2}{9}$$
$$+ (t_1 - \varepsilon)\left(A - t_1 + \frac{2}{3}\varepsilon\right) - (1-\gamma)\,d\left(\frac{4A - 4t_1 + 2\varepsilon}{3}\right), \qquad \text{(A.10)}$$

where $\frac{\partial^2 \tilde{W}_2^{PR_1}}{\partial t_1^2} = -\frac{8}{9}$, i.e., $\tilde{W}_2^{PR_1}$ is concave in t_1.

The remaining cases are the same as in Appendix B.1.

C.2 Proof of Lemma 2

Recall that we assume throughout the paper that $A > d$ and $A > t_i$.

In the text, we state the cases under which country 1 chooses its best response tax level under the three location equilibria. In what follows, we first summarise these cases with the constraint for each case. Second, we compare these constraints to determine which welfare functions are relevant for comparisons. This helps us to divide the values of our parameters into to three ranges as illustrated below. Finally, we rank the welfare levels of country 1 under each range in Lemma 2 (a) to (c). Those lemmas lead to Lemma 2 in the text.

First: the best responses of country 1 are summarised as follows:

Under the TR_2-location equilibrium:

$$\text{if} \begin{cases} t_2 > \gamma d & \hat{t}_1^{TR_2} \rightarrow \hat{W}_1^{TR_2} \text{ unconstrained} \\ t_2 \leq \gamma d & \tilde{t}_1^{TR_2} \rightarrow \tilde{W}_1^{TR_2} \text{ constrained (undercutting)} \end{cases}$$

Under the PR_1-location equilibrium:

$$\text{if} \begin{cases} t_2 \geq 2\gamma d & \check{t}_1^{PR_1} \rightarrow \check{W}_1^{PR_1} \text{ constrained (BCA-constraint violates)} \\ t_2 \geq 2A - 2\gamma d & \ddot{t}_1^{PR_1} \rightarrow \ddot{W}_1^{PR_1} \text{ constrained (NNC violates)} \\ t_2 < 2\gamma d \,\&\, t_2 < 2A - 2\gamma d & \hat{t}_1^{PR_1} \rightarrow \hat{W}_1^{PR_1} \text{ unconstrained} \end{cases}$$

Obviously, there is only one function under the NR-location equilibrium that assumes $t_1 = t_2$.

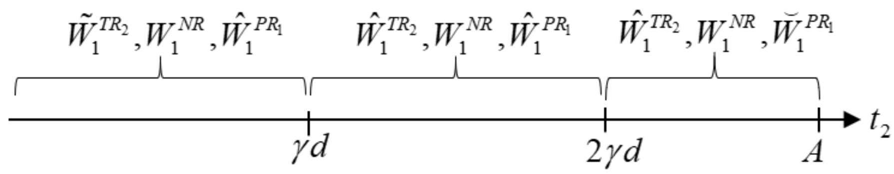
Second, we check whether the two constraints under the PR_1-location equilibrium violate at the same time. We find that if $A > 2\gamma d$, then $A < 2A - 2\gamma d$ and since we assume that $t_2 < A$, thus $t_2 < A < 2A - 2\gamma d$, i.e. the NNC is satisfied. However, the BCA-constraint is not necessarily satisfied since we may have $A > t_2 \geq 2\gamma d$. While if $A \leq 2\gamma d$, we have $2A - 2\gamma d \leq A$, which may lead to $2A - 2\gamma d \leq t_2 < A$, i.e.

the NNC is violated. In such cases, the BCA-constraint is satisfied, $t_2 < A \leq 2\gamma d$. Hence, we either use $\breve{W}_1^{PR_1}$ or $\ddot{W}_1^{PR_1}$ according to the parameter range.

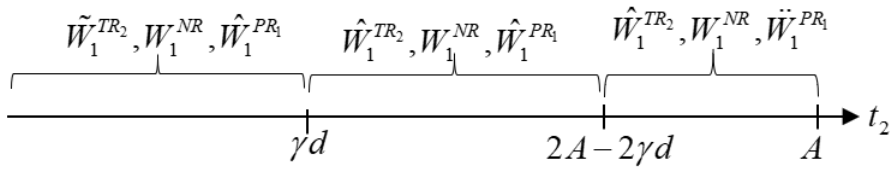
We also need to compare the constraints across the location equilibria. Clearly, $2\gamma d > \gamma d$, hence if $t_2 \geq 2\gamma d$, t_2 is necessarily larger than γd. In such cases, the constrained welfare level $\breve{W}_1^{PR_1}$ is compared with the unconstrained welfare level $\hat{W}_1^{TR_2}$. In addition, we find that $2A - 2\gamma d > \gamma d$ if $A > \frac{3}{2}\gamma d$. As a result, if $t_2 \geq 2A - 2\gamma d$, i.e. if the NNC is violated, we must have $t_2 > \gamma d$. In such cases, the constrained welfare level $\ddot{W}_1^{PR_1}$ is compared with the unconstrained welfare level $\hat{W}_1^{TR_2}$. In contrast, $\gamma d \geq 2A - 2\gamma d$ if $A \leq \frac{3}{2}\gamma d$. This implies that if $2A - 2\gamma d \leq \gamma d < t_2$, we compare the constrained welfare level $\ddot{W}_1^{PR_1}$ with the unconstrained welfare level $\hat{W}_1^{TR_2}$, while if $2A - 2\gamma d \leq t_2 \leq \gamma d$, we compare the constrained welfare level $\ddot{W}_1^{PR_1}$ with the constrained welfare level $\tilde{W}_1^{TR_2}$.

Finally, we divide Lemma 2 into three ranges based on what we have explained above: (a) $A > 2\gamma d$, (b) $\frac{3}{2}\gamma d < A \leq 2\gamma d$ and (c) $d < A \leq \frac{3}{2}\gamma d$. We illustrate the three ranges below and then we derive the proof of Lemma (a) to (c).

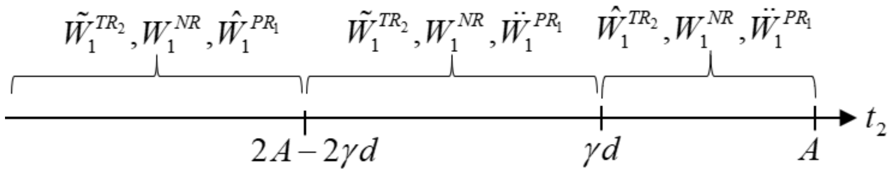
(a) $A > 2\gamma d$



(b) $\frac{3}{2}\gamma d < A \leq 2\gamma d$



(c) $d < A \leq \frac{3}{2}\gamma d$



Lemma. *2(a) $A > 2\gamma d$ (implying $A < 2A - 2\gamma d$)*

i. $2\gamma d \leq t_2 < A$*:* $\hat{W}_1^{TR_2} > W_1^{NR} > \breve{W}_1^{PR_1}$.

ii. $\gamma d < t_2 < 2\gamma d$*:* $\hat{W}_1^{TR_2} > W_1^{NR}$ *and* $\hat{W}_1^{TR_2} > \hat{W}_1^{PR_1}$.

iii. $\underline{t_2} < t_2 \leq \gamma d$*:* $\tilde{W}_1^{TR_2} > W_1^{NR}$ *and* $\tilde{W}_1^{TR_2} > \hat{W}_1^{PR_1}$.

iv. $-\frac{A}{2} < t_2 \leq \underline{t_2}$*:* $\hat{W}_1^{PR_1} \geq \tilde{W}_1^{TR_2} > W_1^{NR}$.

v. $t_2 \leq -\frac{A}{2}$*:* $\hat{W}_1^{PR_1} > W_1^{NR} > \tilde{W}_1^{TR_2}$.

Proof. We use appendices B.1, B.2 and C.1.

(i) It is clear that as long as $\hat{W}_1^{TR_2}$ is feasible, it dominates all other welfare levels. That is, $\hat{W}_1^{TR_2} > \breve{W}_1^{PR_1}$ and $\hat{W}_1^{TR_2} > W_1^{NR}$. Additionally, $\breve{W}_1^{PR_1}$ and W_1^{NR} intersect at two tax levels t_2: A and $-\frac{A}{2}$ for all $\varepsilon \to 0$ such that $W_1^{NR} \leq \breve{W}_1^{PR_1}$for all $t_2 \geq A$, which is not feasible, and for all $t_2 \leq -\frac{A}{2}$, which is not included in this range. Therefore, for $2\gamma d \leq t_2 < A$, we have $\hat{W}_1^{TR_2} > W_1^{NR} > \breve{W}_1^{PR_1}$.

(ii) Since $\hat{W}_1^{TR_2}$ is still feasible, we have $\hat{W}_1^{TR_2} > \hat{W}_1^{PR_1}$ and $\hat{W}_1^{TR_2} > W_1^{NR}$.

(iii) $\tilde{W}_1^{TR_2}$ intersects with $\hat{W}_1^{PR_1}$ at two tax levels t_2: $t_2 = \frac{2}{9}(A + \gamma d \pm \Theta)$ for all $\varepsilon \to 0$ where $\Theta = \sqrt{2(A+\gamma d)(5A - 4\gamma d)}$, such that $\hat{W}_1^{PR_1} \geq \tilde{W}_1^{TR_2}$ for all $t_2 \in \left[\frac{2}{9}(A + \gamma d + \Theta) + \varepsilon, \infty\right)$ and for all $t_2 \in \left(-\infty, \frac{2}{9}(A + \gamma d - \Theta) + \varepsilon\right]$. Tax level $t_2 = \frac{2}{9}(A + \gamma d + \Theta) + \varepsilon$ is larger than γd for all $A > 1.08\gamma d$, and hence it is not included in this range since we have $A > 2\gamma d$. Tax level $t_2 = \frac{2}{9}(A + \gamma d - \Theta) + \varepsilon \leq 0$ for all $A \geq \gamma d$, which holds with strict inequality since we assume $A > d$. In addition, $t_2 = \frac{2}{9}(A + \gamma d - \Theta) + \varepsilon \leq -\frac{A}{2}$ for all $A \in (-\infty, -\gamma d]$ which is not feasible, and $t_2 = \frac{2}{9}(A + \gamma d - \Theta) + \varepsilon \geq -\frac{A}{2}$ for all $A \geq \frac{4}{5}\gamma d$ which holds with strict inequality since we assume $A > d$, hence $t_2 = \frac{2}{9}(A + \gamma d - \Theta) + \varepsilon > -\frac{A}{2}$. We denote this tax level by $\underline{t_2}(\gamma, d, A) = \frac{2}{9}(A + \gamma d - \Theta) + \varepsilon$ in the text. Therefore, we have $\tilde{W}_1^{TR_2} > \hat{W}_1^{PR_1}$ for all $t_2 > \underline{t_2}$. From Appendix B.2, we have $\tilde{W}_1^{TR_2} < W_1^{NR}$ for all $t_1 \leq -\frac{A}{2}$. Therefore, in this range, we have $\tilde{W}_1^{TR_2} > W_1^{NR}$.

(iv) $\hat{W}_1^{PR_1}$ intersects with W_1^{NR} at two levels of t_2: $\frac{2}{11}(A + 3\gamma d) \pm \psi$, where $\psi = \frac{2}{11}\sqrt{6(2A^2 + A\gamma d - 4\gamma^2 d^2)}$, such that $W_1^{NR} \geq \hat{W}_1^{PR_1}$ for all $t_2 \in \left[\frac{2}{11}(A + 3\gamma d) - \psi, \frac{2}{11}(A + 3\gamma d) + \psi\right]$. However, the larger tax level is not included in this range, where $\frac{2}{11}(A + 3\gamma d) + \psi \geq \gamma d$ for all $A > 1.232\gamma d$, which holds since $A > 2\gamma d$, while $\frac{2}{11}(A + 3\gamma d) - \psi$ is larger than $-\frac{A}{2}$ if $A > \frac{\gamma d\left(\sqrt{33}-1\right)}{4}$, which always holds as long as $A > 2\gamma d$, but also if $A > d$ for all $\gamma \leq 0.843$. However, if $\gamma > 0.843$, and if $\frac{\gamma d\left(\sqrt{33}-1\right)}{4} > A > d$, ψ is not defined and we have $W_1^{NR} < \hat{W}_1^{PR_1}$. Applying the same conditions, we find that $\frac{2}{11}(A + 3\gamma d) - \psi$ is also larger than $\underline{t_2}$. Hence, we have $\hat{W}_1^{PR_1} \geq \tilde{W}_1^{TR_2} > W_1^{NR}$. *(v)* For all $t_1 \leq -\frac{A}{2}$, we have $\hat{W}_1^{PR_1} > W_1^{NR} > \tilde{W}_1^{TR_2}$. □

Lemma. *2(b)* $\frac{3}{2}\gamma d < A \leq 2\gamma d$ *(implying* $\gamma d < 2A - 2\gamma d \leq A$*)*

i. $2A - 2\gamma d \leq t_2 < A$*:* $\hat{W}_1^{TR_2} > W_1^{NR} > \ddot{W}_1^{PR_1}$.

ii. $\gamma d < t_2 < 2A - 2\gamma d$*:* $\hat{W}_1^{TR_2} > W_1^{NR}$ *and* $\hat{W}_1^{TR_2} > \hat{W}_1^{PR_1}$.

iii. $\underline{t_2} < t_2 \leq \gamma d$*:* $\tilde{W}_1^{TR_2} > W_1^{NR}$ *and* $\tilde{W}_1^{TR_2} > \hat{W}_1^{PR_1}$.

iv. $-\frac{A}{2} < t_2 \leq \underline{t_2}$*:* $\hat{W}_1^{PR_1} \geq \tilde{W}_1^{TR_2} > W_1^{NR}$.

v. $t_2 \leq -\frac{A}{2}$: $\hat{W}_1^{PR_1} > W_1^{NR} > \tilde{W}_1^{TR_2}$.

Proof. *(i)* W_1^{NR} intersects with $\ddot{W}_1^{PR_1}$ at two levels of t_2, which are A and $-2A + 3\gamma d$ for $\varepsilon \to 0$, such that $\ddot{W}_1^{PR_1} \geq W_1^{NR}$ for all $t_2 \in [A, \infty)$ and for all $t_2 \in (-\infty, -2A + 3\gamma d]$. The first solution is not feasible since it is larger than A, while the second solution is less than $2A - 2\gamma d$ for all $A > \frac{5}{4}\gamma d$, which holds since we assume in this case $A > \frac{3}{2}\gamma d$. Hence, the two intersection points are not included in this range. Therefore, we have $\hat{W}_1^{TR_2} > W_1^{NR} > \ddot{W}_1^{PR_1}$.

(ii) to *(v)* The BCA-constraint and the NNC are satisfied, and, hence we have the same proof as in Lemma 2(a). Note that if $A = 2\gamma d$, $A = 2A - 2\gamma d$, and, hence $t_2 < A = 2\gamma d = 2A - 2\gamma d$, i.e. both constraints are satisfied under the PR_1-location equilibrium and the Lemma starts from (ii). □

Lemma. *2(c)* $d < A \leq \frac{3}{2}\gamma d$ *(implying* $2A - 2\gamma d \leq \gamma d < A$*)*

i. $\gamma d < t_2 < A$: $\hat{W}_1^{TR_2} > W_1^{NR} > \ddot{W}_1^{PR_1}$.

ii. $2A - 2\gamma d \leq t_2 \leq \gamma d$: $\tilde{W}_1^{TR_2} > W_1^{NR}$ *and* $\tilde{W}_1^{TR_2} > \ddot{W}_1^{PR_1}$.

iii. $\underline{t_2} < t_2 < 2A - 2\gamma d$: $\tilde{W}_1^{TR_2} > W_1^{NR}$ *and* $\tilde{W}_1^{TR_2} > \hat{W}_1^{PR_1}$.

iv. $-\frac{A}{2} < t_2 \leq \underline{t_2}$: $\hat{W}_1^{PR_1} \geq \tilde{W}_1^{TR_2} > W_1^{NR}$.

v. $t_2 \leq -\frac{A}{2}$: $\hat{W}_1^{PR_1} > W_1^{NR} > \tilde{W}_1^{TR_2}$.

Proof. *(i)* In this range, the NNC is also violated under PR_1, and, hence, we have the same proof as for the first range in Lemma 2(b). The only difference is that $-2A + 3\gamma d < \gamma d$ for all $A > \gamma d$, which holds.

(ii) In this range, $\tilde{W}_1^{TR_2}$ intersects with $\ddot{W}_1^{PR_1}$ at two levels of t_2: A and $\gamma d - A$ for $\varepsilon \to 0$, such that $\ddot{W}_1^{PR_1} \geq \tilde{W}_1^{TR_2}$ for all $t_2 \in [A, \infty)$ and for all $t_2 \in (-\infty, \gamma d - A]$. However, the first solution is not feasible. In addition, we have $\gamma d - A < 0$ for all $A > \gamma d$, which holds since we have $A > d$. Moreover, this range assumes $t_2 \geq 2A - 2\gamma d$ where $2A - 2\gamma d \geq 0$ for all $A \geq \gamma d$. Hence, the second solution is not included in this range. Therefore, we have $\tilde{W}_1^{TR_2} > \ddot{W}_1^{PR_1}$. Recall that we also have $\tilde{W}_1^{TR_2} > W_1^{NR}$ from Appendix B.2.

(iii) to *(v)* We have the same proof as in Lemma 2(a) and (b). However, we need to mention that in this case in (iii) the tax level $t_2 = \frac{2}{9}(A + \gamma d + \Theta) + \varepsilon$, at which $\tilde{W}_1^{TR_2}$ intersects with $\hat{W}_1^{PR_1}$ is larger than $2A - 2\gamma d$ for all $\frac{4}{5}\gamma d < A < 2\gamma d$, which holds, and, hence this level is not included in this range. □

C.3 Proof of Lemma 3

The NNC requires $\hat{t}_2^{PR_1}(t_1) < A$, which implies $t_1 > 2(1-\gamma)d - A$. The BCA-constraint requires $t_1 > \hat{t}_2^{PR_1}(t_1)$. For this to be true, we must have $t_1 > \bar{t}_1$ with $\bar{t}_1 = \frac{1}{7}A + \frac{6}{7}(1-\gamma)d$. We find that $\bar{t}_1 \geq 2(1-\gamma)d - A$ for all $A \geq (1-\gamma)d$, which generally holds with strict inequality. Thus, we only have to consider the BCA-constraint. We use appendices B.1, B.2 and C.1.

(i) In the range $t_1 > \bar{t}_1$, $\hat{W}_2^{PR_1}$ and W_2^{NR} intersect at two levels of t_1: $t_1 = \frac{1}{11}(A + 14(1-\gamma)d) \pm \Omega$, where $\Omega = \frac{4}{11}\sqrt{4(\gamma-1)^2 d^2 + A(\gamma-1)d - 2A^2}$. However, these two tax levels are not defined for all $A > d$. Hence, we always have $\hat{W}_2^{PR_1} > W_2^{NR}$. From Appendix B.2, we have $W_2^{NR} \leq W_2^{TR_2}$ for all $t_1 \leq -\frac{A}{2}$. In addition, we have $\bar{t}_1 > -\frac{A}{2}$ for all $A > -\frac{4}{3}d(1-\gamma)$, which obviously holds since $A > 0$. Therefore, we have $\hat{W}_2^{PR_1} > W_2^{NR} > W_2^{TR_2}$.

(ii) In the range $t_1 \leq \bar{t}_1$, country 2 chooses its constrained carbon tax $t_1 - \varepsilon$ where $\tilde{W}_2^{PR_1}$ intersects with W_2^{NR} at two tax levels of t_1: A and $-\frac{A}{2}$ for $\varepsilon \to 0$, where $\tilde{W}_2^{PR_1} < W_2^{NR}$ for all $t_1 \leq -\frac{A}{2}$. Therefore, in this range, we have $\tilde{W}_2^{PR_1} > W_2^{NR} > W_2^{TR_2}$. The three curves intersect at $t_1 = -\frac{A}{2}$ for $\varepsilon \to 0$. Strictly speaking, $\tilde{W}_2^{PR_1}$ intersects with W_2^{NR} at $t_1 \approx -\frac{A}{2} + \varepsilon$. (See Appendix B.3 and footnote 11 in the text).

(iii) Finally, in the range $t_1 \leq -\frac{A}{2}$, we have $W_2^{TR_2} \geq W_2^{NR} > \tilde{W}_2^{PR_1}$.

C.4 Proof of Proposition 5

(i) The proof follows directly from Figure 8(b).

(ii) Solving the two response functions as given in (18) and (20) in the text, gives, $t_1^{NE*(P\hat{R}_1)} = \frac{1}{11}A + \frac{1}{11}d(6+2\gamma)$ and $t_2^{NE*(P\hat{R}_1)} = \frac{2}{11}A + \frac{2}{11}d(6-9\gamma)$. A NE exists if $\underline{t_2} \geq t_2^{NE*(P\hat{R}_1)}$, where $\underline{t_2} = \frac{2}{9}(A + \gamma d - \Theta) + \varepsilon$ with Θ as defined in Appendix C.2. This inequality holds if and only if $A \leq \overline{A}_{NE} = \frac{d\left(7\gamma - 12 + 33\sqrt{129\gamma^2 - 136\gamma + 40}\right)}{134}$ for $\varepsilon \to 0$. As also for the threshold $\overline{A}_{NE} > d$ must hold, we find $A \leq \overline{A}_{NE}$ is only feasible if $\gamma > 0.855$. Then, the threshold $\overline{A}_{NE}$ increases in d. Hence, existence of a NE requires d and γ to be large.

(iii) If a NE exits, $t_1^{NE*(P\hat{R}_1)} > 0$ and $t_2^{NE*(P\hat{R}_1)} < 0$ (since a NE exists only if $\underline{t_2} \geq t_2^{NE*(P\hat{R}_1)}$, and, as mentioned in Appendix C.2, $\underline{t_2} < 0$, hence, $t_2^{NE*(P\hat{R}_1)}$ is a subsidy). In addition, $t_2^{NE*(P\hat{R}_1)} \leq -\frac{A}{2}$ if $A \leq d((12\gamma - 8)/5)$. However, this condition violates our assumption $A > d$ for all $\gamma \leq 1.083$. Hence, $t_2^{NE*(P\hat{R}_1)} > -\frac{A}{2}$.

C.5 Proof of Proposition 6

Country 1 chooses between PR_1 and NR. $W_1^{L*}(NR)$ is given in Appendix B.3. In the unconstrained PR_1-location equilibrium, we obtain: $t_1^{L*}(\hat{PR_1}) = \frac{2}{7}A + \frac{1}{7}d(3-2\gamma)$ and $t_2^{F*}(\hat{PR_1}) = \frac{1}{28}A + \frac{1}{7}d\left(\frac{33}{4} - 9\gamma\right)$ from maximising $W_1^{PR_1}\left(t_1, \hat{t}_2^{PR_1}(t_1)\right)$ with respect to t_1, where $\frac{\partial^2 W_1^{PR_1}}{\partial t_1^2} = -\frac{7}{6} < 0$. It is clear that $t_1^{L*}(\hat{PR_1}) > 0$ for all $\gamma \leq 1$. In addition, $t_2^{F*}(\hat{PR_1}) \leq -\frac{A}{2}$ if $A \leq \frac{1}{5}d(12\gamma - 11)$, which violates the NNC $A > d$. Hence, $t_2^{F*}(\hat{PR_1}) > -\frac{A}{2}$. Furthermore, $t_1^{L*}(\hat{PR_1}) > \bar{t}_1$ if $A > d(3-4\gamma)$, which holds as long as $A > d$, and $t_1^{L*}(\hat{PR_1}) < A$ if $A > \frac{1}{5}d(3-2\gamma)$, which also holds as long as $A > d$. Inserting these equilibrium taxes into (17), we obtain $W_1^{L*}(\hat{PR_1})$. We find that $W_1^{L*}(\hat{PR_1}) \geq W_1^{L*}(NR)$ for all $A \leq \overline{A}_1^{PR_1} = \frac{d\left(\sqrt{42}\sqrt{32\gamma^2-8\gamma+9}+52\gamma-15\right)}{17}$ where the threshold $\overline{A}_1^{PR_1}$ increases in d and γ as the term in brackets is larger than zero and increases in γ for all values of γ. In addition, this threshold level is always feasible, i.e., $\overline{A}_1^{PR_1} > d$.

In order to check that country 1 does not prefer to set $-\frac{A}{2} < t_1^L \leq \bar{t}_1$, such that country 2 responds by undercutting t_1, we insert $t_2 = t_1 - \varepsilon$ into (17) and obtain $W_1^{L*}(\widetilde{PR_1})$. We find that $W_1^{L*}(\hat{PR_1}) = W_1^{L*}(\widetilde{PR_1})$ at $\widetilde{t}_1 = \frac{3A^2+(30-48\gamma)Ad+\left(80\gamma^2-72\gamma-9\right)d^2}{28(A-4\gamma d)}$ for $\varepsilon \to 0$. On the one hand, $W_1^{L*}(\hat{PR_1}) \leq W_1^{L*}(\widetilde{PR_1})$ for all $t_1 \in \left[\widetilde{t}_1, \infty\right)$ if $A < 4\gamma d$, which implies that $A < \overline{A}_1^{PR_1}$. Nevertheless, we have $\widetilde{t}_1 > \bar{t}_1$ for all $A < 4\gamma d$. Hence, country 2 would react on its standard reaction function and country 1 chooses $t_1^{L*}(\hat{PR_1})$. On the other hand, we find that $W_1^{L*}(\hat{PR_1}) \leq W_1^{L*}(\widetilde{PR_1})$ for all $t_1 \in \left(-\infty, \widetilde{t}_1\right]$ if $A \geq 4\gamma d$. However, if $A \geq 4\gamma d$, but $A \leq \overline{A}_1^{PR_1}$, we find that $\widetilde{t}_1 \leq -\frac{A}{2}$ for $\varepsilon \to 0$, and hence undercutting is not a best response for country 2. Finally, if $A > \overline{A}_1^{PR_1}$, which implies $A > 4\gamma d$, we have $W_1^{L*}(\hat{PR_1}) \leq W_1^{L*}(\widetilde{PR_1})$ for all $t_1 \in \left(-\infty, \widetilde{t}_1\right]$ where $\widetilde{t}_1 > -\frac{A}{2}$. However, in this case, i.e. if $A > \overline{A}_1^{PR_1}$, $W_1^{L*}(NR) \geq W_1^{L*}(\widetilde{PR_1})$ for all $t_1 \geq -\frac{A}{2}$ for $\varepsilon \to 0$. Therefore, to sum up, for all $A \leq \overline{A}_1^{PR_1}$, we have $W_1^{L*}(\hat{PR_1}) \geq W_1^{L*}(NR)$ and $W_1^{L*}(\hat{PR_1}) > W_1^{L*}(\widetilde{PR_1})$. In contrast, for all $A > \overline{A}_1^{PR_1}$, we have $W_1^{L*}(NR) > W_1^{L*}(\hat{PR_1})$ and $W_1^{L*}(NR) > W_1^{L*}(\widetilde{PR_1})$.

As mentioned in footnote 17, we find that $W_1^{L*}(\hat{TR_1}) \geq W_1^{L*}(\hat{PR_1})$ if $A \leq \frac{d\left(2\sqrt{126\gamma(1-\gamma)+168}-16\gamma-11\right)}{19}$, which violates the NNC for all $\gamma \leq 1$. Hence, we have $W_1^{L*}(\hat{TR_1}) < W_1^{L*}(\hat{PR_1})$. Recall that $W_1^{L*}(\hat{TR_1})$ is the equilibrium welfare level in the unconstrained TR_1-location equilibrium, i.e. $t_2^{F*} = (1-\gamma)d$, (see Appendix B.3).

C.6 Proof of Proposition 7

Country 2 chooses between PR_1 and TR_2. In the unconstrained PR_1-location equilibrium, we obtain $t_1^{F*}(\hat{PR_1}) = \frac{d(2\gamma+9)}{13}$ and $t_2^{L*}(\hat{PR_1}) = \frac{1}{13}d\,(18-22\gamma)$ from maximising $W_2^{PR_1}\left(t_2, \hat{t}_1^{PR_1}(t_2)\right)$ with respect to t_2, where $\frac{\partial^2 W_2^{PR_1}}{\partial t_2^2} = -\frac{13}{18} < 0$. By inserting these tax levels into (19), we obtain $W_2^{L*}(\hat{PR_1})$.

We need first to check under which conditions country 2 could set this tax level. This requires $t_2^{L*}(\hat{PR_1}) < \underline{t_2}$ which holds if $A < \overline{A}_2^{PR_1} = \frac{d\left(11\gamma-9+3\sqrt{185\gamma^2-246\gamma+90}\right)}{13}$ for $\varepsilon \to 0$. However, this condition is feasible only for all $\gamma \geq \widehat{\gamma} = 0.881$. If these conditions do not hold, country 2 sets its tax marginally below $\underline{t_2}$ and country 1 sets its tax at the level in (18), i.e. $t_2^L \lesssim \underline{t_2}$ and $t_1^F = \frac{1}{2}t_2^L + \gamma d$. Inserting these tax levels into (19), we obtain $W_2^{L*}(\underline{PR_1})$. These equilibria under PR_1 need to be compared with the welfare level under TR_2. There are two best responses for country 1 in the TR_2-location equilibrium: $t_1 = \gamma d$ and $t_1 = t_2 - \varepsilon$. Inserting these tax levels into (A.5), we obtain $W_2^{L*}(\hat{TR_2})$ and $W_2^{L*}(\widetilde{TR_2})$, respectively.

First, we find that $W_2^{L*}(\hat{TR_2}) \geq W_2^{L*}(\hat{PR_1})$ if $A \leq \frac{d\left(\sqrt{10764\gamma-7085\gamma^2-3159}-13\gamma\right)}{39}$, which violates our assumption $A > d$. Hence, we always have $W_2^{L*}(\hat{TR_2}) < W_2^{L*}(\hat{PR_1})$. Then, we need to check whether $W_2^{L*}(\hat{TR_2}) \geq W_2^{L*}(\underline{PR_1})$. Due to the complexity of the formula $\underline{t_2}$, it is sometimes not possible to obtain analytical solutions. Thus, we conduct numerical simulations by dividing a large parameter space into four ranges for A and d given the NNC $A > d$ and for all $\gamma \in [0.5, 1]$. We enlarge the set of parameter values in each range as follows: 1) $d \in (0, 50]$, $A \in (d, 100]$, 2) $d \in (0, 100]\,, A \in (d, 500]$, 3) $d \in (0, 500]\,, A \in (d, 1000]$, and 4) $d \in (0, 1000]$, $A \in (d, 10000]$. In addition, and as mentioned by Markusen et al. (1995), we set the marginal cost $c = 0$, since an increase in c and a decrease in a are equivalent (recall $A = a - c$). We find that $W_2^{L*}(\hat{TR_2}) < W_2^{L*}(\underline{PR_1})$ for all the ranges of parameter values. Hence, the two PR_1-location equilibria dominate the unconstrained TR_2-location equilibrium.

Second, we need to consider the constrained TR_2-location equilibrium. We find that $W_2^{L*}(\widetilde{TR_2}) \geq W_2^{L*}(\hat{PR_1})$ if $t_2 \in \left(-\infty, \overline{t_2}\right]$, where $\overline{t_2} = A - 3d\,(1-\gamma) - \frac{1}{26}\sqrt{(7540\gamma^2 - 15288\gamma + 8190)\,d^2 + (3380\gamma - 4056)\,Ad + 1690A^2}$, for $\varepsilon \to 0$. However, $\underline{t_2} \geq \overline{t_2}$ for all $A \geq \frac{4}{5}\gamma d$, which holds with strict inequality since we assume $A > d$. Hence, since $\underline{t_2} \notin \left(-\infty, \overline{t_2}\right]$, country 1 will not respond by undercutting. Therefore, we always have $W_2^{L*}(\hat{PR_1}) > W_2^{L*}(\widetilde{TR_2})$. However, we find $W_2^{L*}(\underline{PR_1}) \leq W_2^{L*}(\widetilde{TR_2})$ if $t_2 \in \left(-\infty, \widetilde{t_2}\right]$ where $\widetilde{t_2} = A - 3d\,(1-\gamma) - \frac{1}{9}\sqrt{13\sqrt{2}\left(\left(\frac{112\gamma-81}{13}\right)d + A\right)\sqrt{A+\gamma d}\sqrt{5A-4\gamma d} + \varpi + \varphi}$ and $\varpi = (729 + 473\gamma^2 - 1134\gamma)\,d^2$ and $\varphi = (280\gamma - 405)\,Ad + 131A^2$ for $\varepsilon \to 0$.

Country 1 responds by undercutting only if $t_2 \geq \underline{t_2}$. We find that $\widetilde{t_2} > \underline{t_2}$ if and only if $A > \overline{A}^{TR_2} = d\left(6 + \frac{1}{2}\left(3\sqrt{\gamma^2 - 8\gamma + 10} - 5\gamma\right)\right)$, where the term in brackets is larger than 1, and, hence this threshold increases in d. However, $\overline{A}^{TR_2}$ decreases in γ for all $\gamma \leq 1$. If $A > \overline{A}^{TR_2}$, i.e., if $\underline{t_2} \in \left(-\infty, \widetilde{t_2}\right]$, the lowest possible tax that country 2 can choose to induce country 1 to undercut its tax is marginally above $\underline{t_2}$, while if $A \leq \overline{A}^{TR_2}$, $\widetilde{t_2} < \underline{t_2}$. Therefore, if $A > \overline{A}^{TR_2}$, $W_2^{L*}(\underline{PR_1}) < W_2^{L*}(\widetilde{TR_2})$ and $t_2^L \gtrsim \underline{t_2}$.

C.7 Proof of Corollary 2

Inserting the Nash equilibrium tax levels under PR_1 in Appendix C.4 into the equilibrium output levels in Appendix C.1 and into (17) and (19), we obtain $E^{NE*}(\hat{PR_1})$, $W^{NE*}(\hat{PR_1})$, $W_1^{NE*}(\hat{PR_1})$ and $W_2^{NE*}(\hat{PR_1})$. From Appendix C.5, we obtain $E^{SE_1^2*}$ and $W^{SE_1^2*}$. Similarly, from Appendix C.6, we get $E^{SE_2^1*}$ and $W^{SE_2^1*}$. Note that we consider only SE_2^1 (a) in Proposition 7.

We start by comparing the equilibrium welfare and emission levels across the two non-cooperative scenarios (i, ii, iii) and then we compare the non-cooperative outcome with the social optimum (iv).

i) If country 1 is the Stackelberg leader, $W_2^{SE_1^2*} \geq W_2^{NE*}$ if $\frac{d(2199-1060\gamma)}{1201} \leq A \leq \frac{d(12\gamma+3)}{5}$, where the upper condition always holds as long as $A \leq \overline{A}_{NE}$, while the lower condition always holds for all $\gamma \geq \check{\gamma} \approx 0.94$ and there is a contradiction for all $\gamma < 0.877$. However, in terms of global welfare, $W^{SE_1^2*} > W^{NE*}$ if $\frac{d(1821-2572\gamma)}{571} < A < \frac{d(12\gamma+3)}{5}$, where both conditions hold as long as $A > d$ and $A \leq \overline{A}_{NE}$. On the other hand, $E^{SE_1^2*} > E^{NE*}$ if $A < \frac{d(12\gamma+3)}{5}$, where this condition always holds since $A \leq \overline{A}_{NE}$.

ii) If country 2 is the Stackelberg leader, $W_1^{SE_2^1*} \geq W_1^{NE*}$ and $W^{SE_2^1*} \geq W^{NE*}$ if $A \leq \frac{d(21-4\gamma)}{13}$, where this condition is satisfied as long as $A < \overline{A}_2^{PR_1}$ and $A \leq \overline{A}_{NE}$ for all $\gamma < \ddot{\gamma} \approx 0.98$. While for $\gamma \geq 0.98$, this condition needs to hold. In addition, $E^{NE*} \geq E^{SE_2^1*}$ if the same condition holds, i.e. if $A \leq \frac{d(21-4\gamma)}{13}$.

iii) We show in i) that global welfare and global emission levels are higher under the leadership of country 1 than under the NE. Hence, we compare now with the leadership of country 2: $W^{SE_1^2*} > W^{SE_2^1*}$ if $\frac{d(5631+1468\gamma)}{4121} - \Psi < A < \frac{d(5631+1468\gamma)}{4121} + \Psi$, where $\Psi = \frac{d(8\gamma-3)336\sqrt{7}}{4121}$. Both conditions always hold as long as $A < \overline{A}_2^{PR_1}$ and $A \leq \overline{A}_{NE}$ and $A > d$. In addition, $E^{SE_1^2*} > E^{SE_2^1*}$ if $A < \frac{d(57+4\gamma)}{39}$, which also always holds as long as $A < \overline{A}_2^{PR_1}$ and $A \leq \overline{A}_{NE}$.

iv) We have $E_S^* \geq E^{NE*}$, if $A \geq \frac{d(15+16\gamma)}{14}$, however, this condition violates the NE existence condition $A \leq \overline{A}_{NE}$. Hence, $E_S^* < E^{NE*}$. If country 1 leads, $E_S^* \geq E^{SE_1^2*}$

if $A \geq \frac{d(39+44\gamma)}{37}$. However, this condition also violates $A \leq \overline{A}_{NE}$, while if a NE does not exists, we have $E_S^* > E^{SE_1^2*}$ if $\frac{d(39+44\gamma)}{37} < A < \overline{A}_1{}^{PR_1}$ (see Proposition 6). If country 2 leads, $E_S^* \geq E^{SE_2^1*}$ if $A \geq \frac{d(12+20\gamma)}{13}$, which also violates $A < \overline{A}_2^{PR_1}$ and $A \leq \overline{A}_{NE}$, and, hence $E_S^* < E^{SE_2^1*}$.

D Appendix of Section 5

D.1 Proof of Corollary 3

(*i*) We show in Appendix C.4 that $t_1^{NE*}(\hat{PR_1}) > 0$ and $t_2^{NE*}(\hat{PR_1}) > -\frac{A}{2} = t_i^{NE*}(NR)$. Hence, global emissions are higher without BCAs. From Appendix B.4 and C.7, we obtain the global welfare level without BCAs, $W^{NE*}(NR)$, and with BCAs, $W^{NE*}(\hat{PR_1})$. We find $W^{NE*}(\hat{PR_1}) \leq W^{NE*}(NR)$ if $A \leq \frac{d(204+244\gamma-\Omega)}{197}$ or if $A \geq \frac{d(204+244\gamma+\Omega)}{197}$, where $\Omega = 66\sqrt{12\gamma - 16\gamma^2 + 47}$. The first solution violates the NNC, while the second solution violates the existence condition of the NE, $A \leq \overline{A}_{NE}$. Hence, $W^{NE*}(\hat{PR_1}) > W^{NE*}(NR)$.

(*ii*) From Appendix B.3, we obtain the equilibrium welfare levels without BCAs, $W_i^{NE*}(NR)$, and from Appendix C.7, we have the equilibrium welfare levels with BCAs, $W_i^{NE*}(\hat{PR_1})$. We find $W_1^{NE*}(NR) \geq W_1^{NE*}(\hat{PR_1})$ if $A \leq \frac{d(488\gamma-120-\xi)}{163}$ or if $A \geq \frac{d(488\gamma-120+\xi)}{163}$ with $\xi = 22\sqrt{6}\sqrt{38\gamma^2 - 8\gamma + 9}$. The first solution violates the NNC, while the second solution violates the threshold level under which the NE exists. Hence, we have $W_1^{NE*}(NR) < W_1^{NE*}(\hat{PR_1})$. Moreover, we find that $W_2^{NE*}(\hat{PR_1}) > W_1^{NE*}(NR)$ for all $\gamma \leq \bar{\gamma} = 0.96$ and for all $0.96 < \gamma \leq 0.976$ if $A < \frac{d\left(976\gamma-768-132\sqrt{49\gamma^2-74\gamma+26}\right)}{95}$, while $W_2^{NE*}(NR) > W_2^{NE*}(\hat{PR_1})\ \forall\ \gamma > 0.976$.

(*iii*) Follows directly from Propositions 2 and 5.

D.2 Proof of Corollary 4

(1) If country 1 is the leader, we have $\hat{A}_1{}^{TR_1} < \overline{A}_1{}^{PR_1}$ from Proposition 3 and 6. This allows us to define three parameter ranges regions: region F with $A \leq \hat{A}_1{}^{TR_1}$, region G with $\hat{A}_1{}^{TR_1} < A \leq \overline{A}_1{}^{PR_1}$ and region H with $\hat{A}_1{}^{TR_1} < \overline{A}_1{}^{PR_1} \leq A$. We use Appendices B.4 and C.7.

i) In region F, $E^{SE_1^2*}(\hat{TR_1}) \leq E^{SE_1^2*}(\hat{PR_1})$ if $A \leq \frac{d(11-12\gamma)}{9}$, which violates the NNC. Hence, global emissions decrease with BCAs. We have shown in Appendix C.5 that country 1 is better off under the PR_1- than the TR_1-location equilibrium, i.e. $W_1^{*SE_1^2}(\hat{TR_1}) < W_1^{*SE_1^2}(\hat{PR_1})$. On the other hand, it is obvious that country

2 becomes worse off with BCAs as it moves away from its best location equilibrium, i.e. unconstrained TR_1-location equilibrium. In terms of global welfare, $W^{SE_1^2*}(\hat{PR_1}) \geq W^{SE_1^2*}(\hat{TR_1})$ if $A \leq \underline{A}^F = \frac{d\left(403-68\gamma+28\sqrt{90\gamma^2+108\gamma-31}\right)}{317}$. This condition holds for $\gamma \leq \tilde{\gamma} \approx 0.83$ because then $\hat{A}_1^{TR_1} < \underline{A}^{WF}$. However, if $\gamma > 0.83$, we need $A < \underline{A}^F$ to hold.

ii) In region G, we showed in Appendix C.5 that the tax levels under the BCA regime are larger than under the No-BCA regime. Thus, global emissions are lower with BCAs, $E^{SE_1^2*}(\hat{PR_1}) < E^{SE_1^1*}(NR)$. It is straightforward to show that country 1 is better off under PR_1. Regarding country 2, we find that $W_2^{*SE_1^1}(NR) > W_2^{*SE_1^2}(\hat{PR_1})$ if $A < \frac{d\left(1468\gamma-1425+42\sqrt{1216\gamma^2-2360\gamma+1145}\right)}{5}$. This condition cannot hold, given the NNC, $A > d$, for all $\gamma < \breve{\gamma} \approx 0.98$, and, thus $W_2^{*SE_1^1}(NR) < W_2^{*SE_1^2}(\hat{PR_1})$. However, if $\gamma \geq 0.98$, this condition must hold as long as $A \leq \overline{A}_1^{PR_1}$, and, hence $W_2^{*SE_1^1}(NR) > W_2^{*SE_1^2}(PR_1)$. Global welfare increases under the BCA regime if $A < \underline{A}^G = \frac{d\left(795+716\gamma+42\sqrt{192\gamma-256\gamma^2+1382}\right)}{709}$. We find that this condition always holds as long as $A \leq \overline{A}_1^{PR_1}$ if $\gamma < \breve{\gamma} \approx 0.98$. However, if $\gamma \geq 0.98$, we need $A < \underline{A}^G$ to hold.

iii) In region H, we have the same outcome with and without BCAs.

(2) If country 2 is the leader, we have $\hat{A}_2^{TR_2} < \overline{A}_2^{TR_2}$ from Proposition 4 and 7. Therefore, we define three parameter regions. Region M with $A \leq \hat{A}_2^{TR_2}$ for all $\gamma < 0.7$, region N with $\hat{A}_2^{TR_2} < A \leq \overline{A}_2^{TR_2}$ and region O with $\overline{A}_2^{TR_2} < A$.

i) In region M, all firms are subject to tax equal to γd under the No-BCA regime. With BCAs, country 2 chooses SE_2^1 in this region (see Proposition 7). Either $t_2^{L*}(\hat{PR_1})$ or $\underline{t_2}$ is a subsidy level. In addition, from (18), $\hat{t}_1^{PR_1} < \gamma d$ since $t_2 < 0$. Hence, both countries set a lower carbon tax, and, consequently, global emissions are higher under the BCA regime. Without BCAs, country 1 achieves the highest welfare level, i.e., $\hat{W}_1^{TR_2}$. Hence, it obviously becomes worse off under the BCA regime. We have shown in Appendix C.6 that $W_2^{L*}(\hat{PR_1}) > W_2^{L*}(\hat{TR_2})$ and $W_2^{*L}(\underline{PR_1}) > W_2^{*L}(\hat{TR_2})$. Hence, country 2 is better off with BCAs. From a global perspective, and if country 2 chooses SE_2^1(a), $W^{SE_2^2*}(\hat{TR_2}) > W^{SE_2^1*}(\hat{PR_1})$ if $A < \frac{6\left(25\gamma^2+173\gamma-108\right)d}{13(46\gamma-27)}$, which must hold as long as $A \leq \hat{A}_2^{TR_2}$. If country 2 chooses SE_2^1(b), we insert $t_2^{L*} \approx \underline{t_2}$ and $t_1^{F*} = \hat{t}_1^{PR_1}$ into (17) to obtain $W_1^{*SE_2^1}(\underline{PR_1})$. Hence, we need to compare $W^{SE_2^2*}(\hat{TR_2})$ with $W^{SE_2^1*}(\underline{PR_1})$. Recall that due to the complexity of $\underline{t_2}$, we sometimes need to resort to numerical simulations. We use the same ranges of parameter d as in Appendix C.6, however, for parameter A, we restrict our simulations in this region to $d < A \leq \hat{A}_2^{TR_2}$ and for only $\gamma \in [0.5, 0.7)$. We find $W^{SE_2^2*}(\hat{TR_2}) > W^{SE_2^1*}(\underline{PR_1})$, and, hence global welfare is higher under the

No-BCA regime in this region.

ii) In region N, firms move from NR to PR_1. In this case, both countries set higher carbon taxes since $\underline{t_2} > -\frac{A}{2}$. Hence, global emissions are lower under the BCA regime. For country 1, we have $W_1^{SE_2^1*}(\hat{PR_1}) \leq W_1^{SE_2^1*}(NR)$ if $A \leq (\geq) \frac{d\left(48\gamma-18-(+)\sqrt{2}\sqrt{548\gamma^2-360\gamma+243}\right)}{13}$, where the first condition violates the NNC, and the second condition violates $A < \overline{A}_2^{PR_1}$. Hence, $W_1^{SE_2^1*}(\hat{PR_1}) > W_1^{SE_2^1*}(NR)$. In addition, $W_1^{SE_2^1*}(\underline{PR_1}) > W_1^{SE_2^1*}(NR)$ for all $A > -\gamma d$, which must hold. We checked this also for the same parameter ranges for d and γ defined in Appendix C.6, while for parameter A, it is restricted to $\overline{A}_2^{PR_1} \leq A \leq \overline{A}_2^{TR_2}$ for all $\gamma \in (0.881, 1]$ and $d < A \leq \overline{A}_2^{TR_2}$ for all $\gamma \in [0.5, 0.881)$. For country 2, $W_2^{SE_2^1*}(NR) < W_2^{SE_2^1*}(\hat{PR_1})$ for all $\gamma \leq 0.96$ and if $A < \frac{d\left(104\gamma-78-3\sqrt{26}\sqrt{40\gamma^2-56\gamma+17}\right)}{13}$ for all $0.96 < \gamma < 0.98$. However, this condition cannot hold for all $\gamma \geq 0.98$ as it would violate the NNC, $A > d$. Using the same ranges of the parameter values as for country 1, we find $W_2^{SE_2^1*}(NR) < W_2^{SE_2^1*}(\underline{PR_1})$ if A is not too large for all $\gamma < 0.964$. However, for all $\gamma \geq 0.964$, $W_2^{SE_2^1*}(NR) > W_2^{SE_2^1*}(\underline{PR_1})$. In terms of global welfare, $W^{SE_2^1*}(NR) \geq W^{SE_2^1*}(\hat{PR_1})$ if $A \geq \frac{d(12+20\gamma+\Lambda)}{13}$ or if $A \leq \frac{d(12+20\gamma-\Lambda)}{13}$ with $\Lambda = 12\sqrt{3\gamma - 4\gamma^2 + 10}$. The first condition violates the threshold $\overline{A}_2^{PR_1}$ and the second condition violates the NNC. Hence, $W^{SE_2^1*}(\hat{PR_1}) > W^{SE_2^1*}(NR)$. In addition, we find $W^{SE_2^1*}(\underline{PR_1}) > W^{SE_2^1*}(NR)$ for the range of parameter values in this region.

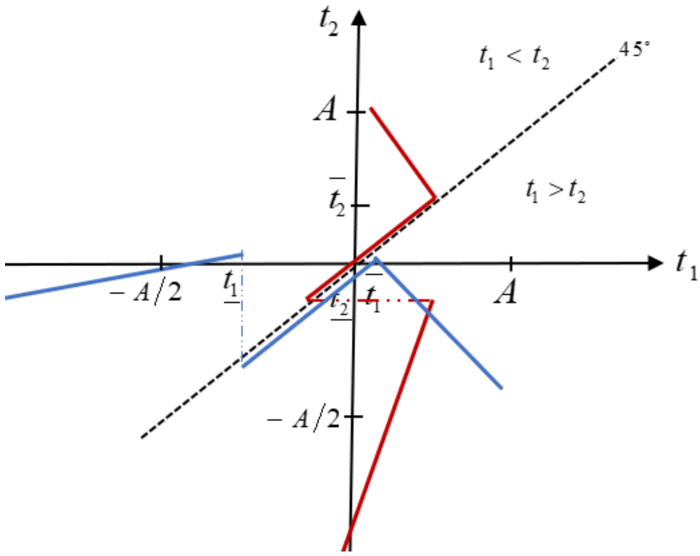
iii) In region O, equilibrium taxes under the BCA regime are $t_2^{*L} \gtrsim t_1^{*F} \gtrsim \underline{t_2}$, which are larger than $-\frac{A}{2}$ under the No-BCA regime. Hence, global emissions are lower with BCAs. Inserting $t_2 = t_2^{*L} \gtrsim \underline{t_2}$ into (A.3) we obtain $W_1^{SE_2^2*}(\tilde{TR_2})$, and inserting $t_1^{F*} \gtrsim \underline{t_2}$ into (A.5), we obtain $W_2^{SE_2^2*}(\tilde{TR_2})$. We find $W_1^{SE_2^2*}(\tilde{TR_2}) < W_1^{SE_2^1*}(NR)$ if $A \leq -\gamma d$, which violates the NNC. In addition, we checked that $W_1^{SE_2^2*}(\tilde{TR_2}) > W_1^{SE_2^2*}(NR)$ for the parameter values of d and γ as defined in Appendix C.6, while for parameter A we assume $A > \overline{A}_2^{TR_2}$. For country 2, we find $W_2^{SE_2^1*}(NR) > W_2^{SE_2^2*}(\tilde{TR_2})$ if $A > \frac{d\left(492-472\gamma-24\sqrt{2}\sqrt{15\gamma^2-37\gamma+20}\right)}{169}$ for all $\gamma < 0.79$ and if $A > \frac{4}{5}\gamma d$ for all $\gamma \geq 0.79$. The two conditions must hold as long as $A > d$ and $A > \overline{A}_2^{TR_2}$. Finally, from a global perspective, we find $W^{SE_2^2*}(\tilde{TR_2}) \geq W^{SE_2^1*}(NR)$ if $\frac{4}{5}\gamma d \leq A \leq \underline{A}^O = 2d\left(12\sqrt{10-\gamma} - 2\gamma + 39\right)$. Therefore, $W^{SE_2^2*}(\tilde{TR_2}) < W^{SE_2^1*}(NR)$ if $A > \underline{A}^O$.

E Bilateral BCA-policy

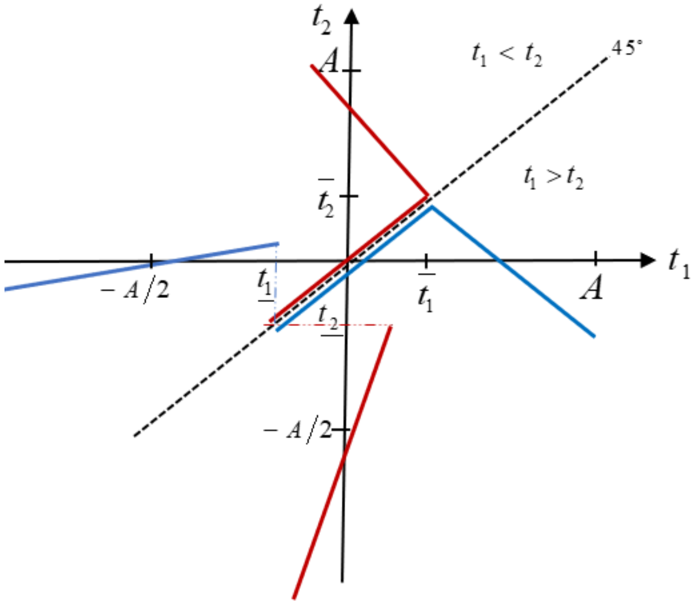
Consider the case in which both countries could impose BCAs, country i if $t_i > t_j$ and country j if $t_i < t_j$. For bilateral BCAs, there are three possible location equilibria: NR, PR_1 and PR_2. The first two equilibria are covered in the main text. In the PR_2-location equilibrium, i.e., if $t_2 > t_1$, only the plant of firm 2 that supplies the home market remains in country 2, while the other plant that supplies country 1 will relocate.

All the analysis that we have done for country 1 in PR_1-location equilibrium will be the same for country 2 in the PR_2-location equilibrium if $t_2 > t_1$. Therefore, also the best response function of country 2 is discontinuous in this case. Hence, in the simultaneous game, a NE does not always exist as shown in Figure A.1 below.

In case of symmetric countries, no NE exists, as this has been derived in the main text for a unilateral BCA-policy. If a NE exists, it would be partial relocation of firm 1, i.e., $t_1 > t_2$. The intersection of the standard best response functions is not possible in the PR_2-location equilibrium. That is, the condition $\underline{t_1} \geq t_1^{*NE}(PR_2)$ is not feasible where the critical tax level for country 2 is $\underline{t_1} = \frac{2}{19}\left(2A + 3(1-\gamma)d - \sqrt{12(1-\gamma)d(A-4) + 42A^2}\right) + \varepsilon$. In addition, we have $t_1^{*NE}(PR_2) > t_2^{*NE}(PR_2)$ for all $\gamma > 0.66$. Therefore, in equilibrium, only country 1 will impose BCAs. Existence of a NE in this case requires similar conditions as derived in Section 4. That is, global marginal damages and asymmetry among countries need to be sufficiently high, i.e., we need $\underline{t_2} \geq t_2^{*NE}(PR_1)$ as derived in the main text. However, now, $\underline{t_2} = \frac{2}{19}\left(2A + 3\gamma d - \sqrt{12\gamma d(A-4) + 42A^2}\right) + \varepsilon$, which is larger than the critical tax level if only country 1 imposes BCAs. Therefore, in the case of bilateral BCAs, the range of parameter values for which a NE exists would be larger than in the case of unilateral BCAs, as considered in the main text. It is obvious that $\underline{t_1} \leq \underline{t_2}$ if $\gamma \geq 0.5$. The complexity of the analysis if country 2 is the Stackelberg leader would now also extend to Stackelberg leadership of country 1 under a bilateral BCA regime.



(a) Existence of a NE 'asymmetric countries'



(b) Non-existence of a NE 'symmetric countries'

Figure A.1: Nash Equilibrium with Bilateral BCAs

Part VI

Summary of Conclusions

1 Summary of Results

In this thesis, we studied the decisions of two asymmetric countries regarding their actions to mitigate climate change using a game-theoretic analysis. Essay 1 presented the basic features of the interaction among countries in mitigating climate change. Generally, countries decide on their climate action based on trading-off the benefits and costs of curbing greenhouse gas (GHG) emissions. However, those benefits and costs differ among countries to such extent that some countries could even be worse off under a cooperative outcome than under no cooperation, and hence cooperation might not be individually rational. We demonstrated that asymmetries across countries greatly affect the outcomes of climate change negotiations and the allocation of abatement burdens. In Essays 2, 3 and 4, those benefits and costs were introduced explicitly by employing an extended version of a strategic trade model due to Brander and Spencer (1985). In these three essays, countries choose their climate policy level, which does not only affect their environmental damages, but also has impacts on governmental income, consumers and the profits of firms that compete in outputs, i.e. in a Cournot fashion.

International trade and the movement of firms raise concerns among nations about carbon leakage and the competitiveness of their firms resulting from differential national climate policies. Therefore, governments may be inclined to implement inefficiently lax environmental policies to protect their home firms. A recently proposed solution to those concerns is to offset the differences between national carbon prices using border carbon adjustment (BCA) measures such as carbon tariffs and export rebates. Implementing BCAs can be considered as a switch to design climate policies based on consumption rather than production. Our study considered different forms of BCAs: a) carbon tariffs, which fully adjust the difference between national carbon taxes, b) carbon tariffs are complemented with export rebates where the export rebate rate is chosen optimally, c) carbon tariffs combined with a full export rebate, which is de facto a unilateral consumption-based carbon tax.

Essay 2 showed that BCAs can partially or completely correct some distortions affecting the choice of carbon taxes, in particular the profit-shifting and carbon leakage effects. As a result, countries may adjust their carbon tax upward. We showed that carbon tariffs with a full export rebate could restore the effectiveness of the environmentally more concerned country's carbon tax by fully internalising its own damages. Nevertheless, tax levels would be set below individual marginal damages due to the optimal response of the other country. Only a bilateral consumption-based carbon tax could be set equal or even above individual marginal damages. We also showed that either unilateral or bilateral consumption-based taxes could surpass the socially optimal policy level. However, this could occur if the environmental damage is not serious.

Essay 3 took a closer look both at the impacts of these measures on individual and global welfare, and also on their strategic role to enforce cooperation. On the one hand, we found that a gradual switch by the environmentally more concerned country to a unilateral consumption-based tax, using BCA-measures, constitutes escalating penalties to enforce full cooperation, with a uniform socially optimal tax. In addition, escalating penalties is credible in the sense that the full set of BCA-threats will be exploited to establish full cooperation. On the other hand, escalating penalties are global welfare distorting should they be implemented compared to the social optimum. We conclude that BCAs can be successful in enforcing cooperation, but whenever the potential gains from full cooperation are expected to be rather significant, even the harshest punishment fails to establish cooperation.

Essay 4 studied the impacts of introducing BCAs (carbon tariffs) if firms can completely close down and relocate abroad their plants. Without BCAs, profits of firms are based on the location of production, and consequently each firm will locate with all its plants in a country which sets a lower carbon tax. If countries choose their carbon taxes simultaneously, they result in a 'race to the bottom' with high emission subsidies. This negative outcome can be avoided if governments move sequentially and the Stackelberg leader chooses the 'wise chicken' equilibrium, which leads to higher welfare levels and lower global emissions. This equilibrium implies that all firms locate in the follower's country and face a higher carbon tax than under the race-to-the-bottom equilibrium. With BCAs, the location of production becomes profit neutral for the plants supplying the country imposing carbon tariffs. This additional measure allows the country to set a higher carbon tax without losing all its plants and gain a strategic advantage by shifting tax revenues from abroad to home. We have thus showed that BCAs reduce the pressure on the race-to-the-bottom if taxes are chosen simultaneously or sequentially.

Comparing a strategic trade model with fixed versus endogenous plant location offers an opportunity to meaningful observations, firstly, under a regime without BCAs. In a fixed plant location model, a) the reaction functions of countries are downward sloping, implying that carbon taxes are strategic substitutes. b) Nash equilibrium taxes would be sub-optimal due to the profit-shifting incentive but there would be no race-to-the-bottom. c) Equilibrium taxes would be asymmetric as long as marginal damages of countries are different. In contrast, in an endogenous plant location model, a) the reaction functions of countries have three segments, one of which is upward sloping, implying that carbon taxes are strategic complements. b) Governments have incentives to not only shift rents, but attract total profits of foreign firms which lead to the Nash equilibrium with a race-to-the-bottom. c) The Nash equilibrium policy level is a symmetric subsidy, irrespective of the asymmetry among countries in terms of their damage evaluations. Secondly, under a regime with BCAs (carbon tariffs) reveals that a Nash equilibrium always exists in a model with fixed plant location leading to higher carbon taxes in both countries and lower global emissions than in one without BCAs. However, in a model with endogenous plant location, a Nash equilibrium may not exist due to the discontinuity of reaction functions. Nevertheless, if an equilibrium exists, BCAs also lead to higher carbon taxes in both countries than in the ones without BCAs. This result is also confirmed in a Stackleberg equilibrium. Hence, taken together, our results showed that BCAs can support more ambitious climate policies under the more general assumption of endogenous plant location.

2 Discussion and Possible Future Research

Overall, conclusions drawn by this thesis provide useful insights into designing climate policies. In a strategic context, carbon leakage and competitiveness concerns can be eliminated if countries impose a uniform global carbon tax or multilateral consumption-based carbon taxes. However, both solutions need coordination among countries which might be difficult to agree upon, at least in the short run. Therefore, unilateral BCA-measures are more likely to emerge in the near future. Although these measures could lead to more stringent climate policies, they should be considered with cautions. Countries with relatively higher carbon price, typically environmentally more concerned countries, would be better off implementing these measures, and in contrast countries on which these measures are imposed might be worse off, which could result in a global welfare loss. Therefore, BCA-measures should ideally be used in the first place as a threat to enforce climate agreements.

Across BCAs measures, adding export rebates to carbon tariffs would result in higher global emissions and lower global welfare than restricting BCAs to imports only. Therefore, our results support the argument that export rebate will less likely be defended on an environmental basis, and also in terms of global welfare. Moreover, we showed that the strategic role of BCAs to support a stricter non-cooperative policy level in the country confronting these measures is triggered by carbon tariffs, whereas export rebates weaken this role thus questioning the effectiveness of export rebates for the internalisation of global damages.

In this thesis, two assumptions allow us to analyse the impacts of BCAs that have been designed in accordance with the World Trade Organization (WTO) rules. More specifically, we consider a two-country model and assume an identical emission/carbon intensity across firms. Therefore, according to the national treatment principle (Article III) of the General Agreement on Tariffs and Trade (GATT), we assume that BCAs, either on imports or exports, are imposed in such a way that foreign producers do not face higher effective carbon payments than domestic producers. In practice, the implementation of unilateral BCA-measures would confront certain legal and practical challenges. For instance, generalising our assumptions to an n-country model with different carbon intensities and/or different carbon prices would imply an unequal treatment between domestic and foreign producers and among foreign producers. BCAs would also not be imposed against countries with comparable or higher carbon prices. In such cases, BCA-measures will violate not only Article III, but also the Most-Favoured-Nation (MFN) principle (Article I) which requires all imported 'like' products to be subject to the same treatment.

Given the potential legal challenges, there is a trade-off between an effective and a legal design of BCAs. Regarding import side, an effective BCA-measure requires that carbon tariffs fully adjust the difference between domestic and foreign carbon taxes. This regulation implies that tariffs would be imposed at different rates depending on the carbon tax in the exporting country, with the highest effective tariff rates imposed against countries with the lowest carbon tax. In such cases, as shown in the thesis, BCAs would incentivise each exporting country to raise its carbon policy level to avoid tariffs, and will both mitigate carbon leakage and protect domestic plants from relocation to less regulated countries. Similarly, in order to reduce foreign emissions effectively, carbon tariffs should be calculated based on the carbon intensity of the foreign production, where countries with the highest carbon intensity would face the highest tariff rate. This would not only protect domestic plants from relocation, but could even attract foreign plants to low-carbon intensity regions, which in turn might encourage countries to adopt clean technologies. Nevertheless, irrespective of the legal issues, information about the carbon intensity of the foreign production process might be difficult and costly to obtain. With respect to export rebates, these should be based on the carbon intensity of the domestic production.

If BCAs were designed to meet Article I and III of the GATT, all imported products should face the same carbon tariffs. Therefore, the government which imposes BCAs could set a uniform tariff rate that would possibly adjust the difference between the domestic carbon price and the highest carbon price (but still below the domestic one) among the exporting countries or the average carbon price in exporting countries. This would greatly undermine the effectiveness of BCAs to tackle carbon leakage. In addition, the strategic effect of carbon tariffs on certain exporting countries would be weakened in particular if there is a wide gap between the carbon prices of exporting countries. Similarly, it has been suggested that BCAs based on carbon intensity from the best available technology or the carbon intensity of the importing country could be legal under the WTO (Ismer and Neuhoff, 2007; Mattoo et al., 2009). However, in both cases, the effectiveness of BCAs to control foreign emissions would dwindle. Regarding BCAs on exports, rebates in excess of a full rebate could be considered as an export subsidy and hence might not be allowed according to the Subsidies and Countervailing Measures Agreement. However, our results show that giving a full rebate or full adjustment is not always in the best interest of governments.

In order to implement BCAs with an effective environmental outcome, these measures could pass the legal issues above under the umbrella of Article XX. The article provides general exceptions for the adoption of certain policy measures which are not compatible with WTO rules, but necessary to protect the environment. However, applying these measures should follow a justifiable discrimination against certain countries or between domestic and foreign producers. These measures, furthermore, should not restrict international trade or constitute a disguised protection.[1] Our results show that BCAs on imports allow countries to impose higher carbon taxes and hence reduce global emissions. Therefore, adjustments on imports could potentially be legal based on Article XX. In addition, it could be justified that discrimination might mitigate carbon leakage rather effectively and encourage other countries to lower their emissions. To avoid any protectionist motive for using BCAs, they should only be implemented as a second best option if a global climate agreement could not be reached. In addition, compensations to countries that might suffer a welfare loss from facing BCAs could be considered through allocating the tariff revenues to technology transfers or the Green Climate Fund to support their mitigation and adaptation efforts (Böhringer et al., 2012; Sakai and Barrett, 2016). We have shown that although adding export rebates is a stronger policy to control carbon leakage than carbon tariffs alone, it would be less likely to be legitimate under Article XX.

[1]See Article XX in GATT Analytical Index (pre-1995).

Multilateral full BCAs (that is multilateral consumption-based carbon taxes) could be recommended as a legal and effective measure. This would mean that foreign producers receive a full export rebate in the home country and pay carbon tariffs equivalent to the carbon tax paid by domestic producers in the foreign (destination) country. In such cases, all imported products are subject to the same tariff rate with an equal treatment between the domestic and imported products. However, as mentioned above, this regime needs coordination between countries to avoid double taxation.

Most of the discussion about BCAs relates to carbon taxes which could, arguably, qualify as an indirect tax, and ones that are adjustable at the border either on imports or exports (given compatibility with Article III). Extending BCAs to other climate policies such as cap-and-trade implies that foreign producers would be required to purchase emissions allowances on terms comparable to domestic producers, and payments to buy allowances would be rebated to exporters. However, the free allocation of allowances would be less likely to justify the application of BCAs, and hence, as suggested by many scholars, such as Pauwelyn (2007), Monjon and Quirion (2010) and Fischer and Fox (2012), the auctioning of allowances might be necessary to impose BCAs, in particular if they are allowed under Article XX.

Our work can be extended in several directions for future research. First, we considered two-country models which could be extended to an n-player asymmetric agreement formation game. Second, different aspects of asymmetry can be assumed, for instance carbon intensity, and also different anti-leakage measures like output-based rebating. Third, further studies could investigate under which conditions countries would agree to switch to a multilateral consumption-based carbon pricing approach. Fourth, for the endogenous plant location model, further research could consider other measures such as export rebates and compare production-based with consumption-based taxes in this set-up.

References

Böhringer, C., Balistreri, E. J., and Rutherford, T. F. (2012). The role of border carbon adjustment in unilateral climate policy: overview of an energy modeling forum study EMF 29. *Energy Economics*, 34:S97–S110.

Brander, J. A. and Spencer, B. J. (1985). Export subsidies and international market share rivalry. *Journal of International Economics*, 18(1-2):83–100.

Fischer, C. and Fox, A. K. (2012). Comparing policies to combat emissions leakage: border carbon adjustments versus rebates. *Journal of Environmental Economics and Management*, 64(2):199–216.

Ismer, R. and Neuhoff, K. (2007). Border tax adjustment: a feasible way to support stringent emission trading. *European Journal of Law and Economics*, 24(2):137–164.

Mattoo, A., Subramanian, A., Van Der Mensbrugghe, D., and He, J. (2009). Reconciling climate change and trade policy. The World Bank.

Monjon, S. and Quirion, P. (2010). How to design a border adjustment for the European Union Emissions Trading System? *Energy Policy*, 38(9):5199–5207.

Pauwelyn, J. (2007). U.S. federal climate policy and competitiveness concerns: the limits and options of international trade law. Working Paper 07-02, Nicholas Institute for Environmental Policy Solutions, Duke University.

Sakai, M. and Barrett, J. (2016). Border carbon adjustments: Addressing emissions embodied in trade. *Energy Policy*, 92:102–110.